PERGAMON

Vision Research 39 (1999) 231–256

Vision Research

Neural recoding in human pattern vision: model and mechanisms

Lynn A. Olzak *, James P. Thomas

Department of Psychology, University of California, Los Angeles, CA 90095-1563, USA

Received 10 June 1996; received in revised form 26 March 1998

Abstract

We describe a model of neural recoding in spatial vision that specifies how the outputs of selected units akin to V1 cells are normalized and combined to signal information about particular stimulus attributes. The recoding portion of the model is linked to psychophysical behavior via a two-stage signal-detection decision module that specifies how the outputs of the combining mechanisms are used in making fine spatial discriminations. We describe how masking and cue summation experiments isolate each of the processing stages, how earlier results from such studies guided development of the model, and we demonstrate how these procedures permit empirical estimates of model parameters as well as tests of alternative formulations. An important part of our work describes the characteristics of two complementary types of higher-level mechanisms isolated from previously published discrimination data. One sums normalized primary-level responses over disparate frequencies to signal precise information about the orientation of a stimulus; the other sums over all orientations to signal the spatial grain of texture-like patterns. We demonstrate how the model accounts for a large body of previously published discrimination data, and present the results of a new quantitative test of model predictions.

Keywords: Discrimination; Masking; Higher-order mechanisms; Orientation; Spatial frequency

1. Introduction

It is generally believed that the processing of visual pattern information follows a hierarchical structure, proceeding along multiple, parallel pathways in the brain. Recoding initially begins in the retina, and proceeds through the LGN to primary visual cortex. Cortical processing is thought to begin with an elementary representation of local spatial frequency, orientation, and contrast, presumably jointly represented by responses of simple cells in primary visual cortex (V1). Much is known about the initial analysis process and the representation of visual information in V1. Much less is known about how the early neural representations are transformed at higher levels of processing to signal information needed to perform visual tasks on complex, real-world objects.

Neuroscientists have identified multiple processing streams that proceed through a number of anatomically defined layers. At least one stream is thought to be specialized for pattern or form vision. This stream proceeds from V1 through V2, V4 and inferotemporal cortex (IT) (Mishkin, Ungerleider & Macko, 1983; Van Essen & Maunsell, 1983; DeYoe & Van Essen, 1988). Physiological evidence is mounting that cells at each ensuing layer respond to more and more complex stimuli, ranging from those that respond to luminance edges and illusory contours at V2 (Peterhans & von der Heydt, 1991) to others that prefer rather complex non-cartesian forms at V4 (Gallant, Braun & Van Essen, 1993), and to very complex forms such as stars etc. at IT (Gross, Bender & Rocha-Miranda, 1969; Tanaka, 1993). The neural processes underlying the changes in selectivity from layer to layer are not yet known, however; nor do we understand the functional significance of these transformations for visually guided behavior.

In this paper, we present a model of how responses from units at initial cortical layers of processing are transformed and used to distinguish among complex patterns that differ only very slightly from one another in orientation, spatial frequency, or contrast. The differences are small enough that performance falls into the hyperacuity (Westheimer, 1975) range. The model was developed to account for a series of discrimination results we reported in previous work (Thomas & Olzak,

* Corresponding author. Tel.: +1 513 2591754; fax: +1 513 2592420; e-mail: olzakla@muohio.edu.

0042-6989/98/$ - see front matter
PII: S0042-6989(98)00122-9

1990; Olzak & Thomas, 1991, 1992) and provides a framework within which to interpret those results and explore new phenomena.

In Section 2 of this paper, we provide a brief background describing our motivation in developing the model. Section 3 provides a formal description of the model. In Section 4, we describe in more detail our earlier discrimination experiments. In that section, we focus on patterns of results found within experiments, and describe how results of masking and cue-summation experiments led us to propose that both broadly-tuned gain control processes and higher-level summing circuits intervene between the initial cortical representations and psychophysical decision processes. Section 5 focuses on the pattern of masking and cue-summation results across the series of 12 earlier experiments. In that section we show how these results, in the context of the model, suggest that fine pattern discriminations are mediated by a small number of higher-level summing circuits, each specialized to signal information about the spatial grain, orientation, or contrast of a pattern. We provide a test of the model in Section 6, and discuss the adequacy of the model to account for old and new data in the final section of the paper, Section 7.

The model we describe follows a general three-stage hierarchical processing structure typical of many current models of spatial vision. The stages include a primary layer of linear tuned filters, a nonlinear processing stage that contains both a within-pathway nonlinearity and a pooled gain control process, and a higher-level summing stage that selectively combines normalized responses of units tuned to different ranges of spatial frequency and orientation. Although our model shares some features in common with other current models at the first two levels of processing, properties of the higher-level summing mechanisms we have inferred from our data differ in fundamental ways from other conceptualizations.

The higher-order summing mechanisms we have isolated comprise at least two different families that appear suited to signal abstract information about different aspects of a stimulus. These families appear to represent independent processing streams (Olzak & Wickens, 1997), each specialized to process a different type of information. One family is specialized to provide information about the orientation of a stimulus feature, such as an edge or line, and may be a general mechanism that signals information about the orientation of object borders. It may also signal the dominant orientation of a surface texture. The other family is specialized to provide information about the textural grain of a surface pattern, and may carry information about the contrast of a pattern, as well.

In the context of the model, the higher-order summing units we describe are differentiated from one another by the particular sets of primary-level pathways over which they sum. The orientation-signalling units selectively sum over wide ranges of spatial frequency but only within limited orientation bands. The textural-grain mechanisms sum over units tuned to all orientations but only within limited frequency bands. Contrast information is carried via pathways that have similar higher-level summing characteristics to that signalling spatial frequency, but that differ in other respects (Thomas & Olzak, 1997).

A central question is whether the mechanisms we describe are general and play a role in tasks other than fine spatial discriminations. Thus, it is important to place our model and mechanisms in the context of other current work in pattern perception. A substantial number of psychophysical models have recently been developed in different domains to characterize pattern processing stages beyond (or perhaps in parallel with) the initial V1 simple-cell representation. Although current models differ significantly from one another in purpose, concept, and in quantitative detail, certain common themes emerge.

An important precursor to many current models, developed by Klein, Stromeyer and Ganz (1974) to account for apparent frequency shifts following adaptation, contained many of the structural elements now incorporated into more formal quantitative models of higher level processing. Specifically, they suggested that the mechanisms underlying perceptual judgments were organized hierarchically and that different visual tasks (e.g. detection as opposed to perceived frequency) tapped different levels in the hierarchy. Klein et al. also suggested the concept of higher-level integrators, which in their model were conceptualized as nonlinear inhibitory interactions occurring over space, but limited to mechanisms with similar tuning characteristics in the Fourier domain.

Although the idea of hierarchical processing in vision was not unique, its application to the field of spatial vision in 1974 was quite innovative. Many, if not most, current models of pattern vision now include some form of hierarchical processing, often displaying a characteristic three-stage structure. This structure typically begins with a linear filtering stage, is followed by a nonlinear processing stage, and culminates with a second stage of filtering in which responses of a subset of the nonlinearly transformed units are 'collected' or summed in some fashion. A plethora of such models have been developed over the past decade to account for different empirical findings. The sheer number may reflect the notion put forth by Klein and his colleagues that different tasks tap different mechanisms at different levels of processing. On the other hand, it seems unlikely that every visual task requires positing a different set of nonlinear and 'higher-level' processes and mechanisms (therefore requiring different models). The

surprising similarity among many current models of mechanisms subserving different tasks suggests the possibility that different tasks are tapping not different mechanisms, but instead are revealing different aspects of a limited number of processes or mechanisms.

Graham and Sutter (1998) summarize many current models of this form in the introductory section of their recent paper, and classify current multistage conceptualizations into two major categories: *complex channels* and *higher-order mechanisms*. As Graham and Sutter have defined them, complex channels are those linear-nonlinear-linear multistage processes which, at the second filtering stage, sum or collect information from receptive fields of similar shape but which differ in spatial position. This class includes mechanisms that have been variously termed non-Fourier mechanisms, second-order mechanisms, collector units and collator units.

Higher-order mechanisms, on the other hand, are defined as those which, at the stage of processing that follows the nonlinearity, sum or collect information from the same position, but over receptive fields that differ in shape. Higher-order units include edge detectors, mechanisms that mediate certain types of texture discriminations, and the mechanisms that we describe in the current paper. Graham and Sutter point out that it is not yet clear to what extent these two types of mechanisms overlap, share processing stages, or are otherwise related.

In order to avoid as much confusion as possible at the outset, we use the Graham-Sutter classification scheme to place our model in the context of others. However, we note that our model might be considered a hybrid of their two categories. The current description of our higher-level mechanisms is based on empirical work performed with complex patterns whose overlapping components differed greatly in spatial frequency and/or orientation. This work firmly places our mechanisms in the 'higher-order' class. However, we fully expect that the receptive fields of these mechanisms have spatial profiles in addition to the characteristic profiles we describe in the Fourier domain. Thus, it is not yet clear in some cases how the mechanisms we describe relate to complex channels.

Models and mechanisms that clearly fall into the complex channel category include recent work on mechanisms mediating motion perception and some findings in texture processing (Adelson & Bergen, 1985; Bergen & Adelson, 1986; Bovik, Clark & Geisler, 1990; Chubb & Landy, 1991; Sagi, 1991; Graham, Beck & Sutter, 1992; Wilson & Richards, 1992; Werkhoven, Sperling & Chubb, 1993; Chubb, McGowan, Sperling & Werkhoven, 1994; Derrington & Henning, 1994; Solomon & Sperling, 1994, 1995; Sperling, Chubb, Solomon & Lu, 1994; Wilson, 1994; Wilson & Kim, 1994; Graham & Sutter, 1996, 1998). Although preliminary evidence we have reported suggests that the mechanisms we describe may also play some role in texture segregation tasks (Stankiewicz, Thomas & Olzak, 1995; Laurinen, Olzak & Peromaa, 1997), the conceptualizations at present seem fundamentally different between this class of complex channel models and ours.

It is less clear to us whether collator or collector units, which Graham and Sutter place in the complex-channels category, are fundamentally different from the orientation-signalling mechanisms we have isolated in our discrimination studies. These terms have been widely used to describe second-order units with large receptive fields that collect information along a path or aligned information over gaps in space. Similar units are currently presumed to mediate the perception of illusory contours, to underlie contour integration (Morgan & Hotopf, 1989; Field, Hayes & Hess, 1993; Moulden, 1994; Hess & Field, 1995), to play a role in findings of long-range facilitation and inhibition (Polat & Norcia, 1996) and to mediate some large-scale localization tasks (Burbeck & Yap, 1990a,b; Burbeck, 1991; Hess & Holliday, 1992; Hess & Hayes, 1993, 1994; Wang & Levi, 1994; Hess & Badcock, 1995; Waugh & Levi, 1995; Levi & Waugh, 1996; Mussap & Levi, 1997). Although we have not yet determined the spatial receptive field size of the orientation signalling mechanisms we describe, our mechanisms share some important defining properties with previously described collator or collector units. We return to this point in the Discussion.

Among models that seem to fall into the higher-order class as defined by Graham and Sutter, perhaps the most closely related to that proposed in present paper are those of mechanisms underlying edge detection and edge localization (Marr & Hildreth, 1980; Watt & Morgan, 1984, 1985; Morrone & Burr, 1988; Georgeson, 1992; Burr & Morrone, 1994), phase discrimination (Meese, 1995), complex pattern perception (Derrington & Badcock, 1985; Georgeson, 1994; Meese & Georgeson, 1996a,b), and possibly stereoacuity (Wilcox & Hess, 1996). Although the higher-level summing units we describe differ from other conceptualizations, we note that the empirical methods we have developed permit us to test and distinguish among key assumptions that differentiate the various models.

The processes we describe in the nonlinear stage of our model grow directly out of earlier work in contrast discrimination and/or perceived contrast (Smith & Thomas, 1989; Albrecht & Geisler, 1991; Cannon & Fullenkamp, 1991, 1996; Heeger, 1991, 1993; Foley, 1994), and rely on physiological evidence for both within-pathway nonlinearities (Albrecht & Hamilton, 1982; Albrecht & Geisler, 1991) and for gain control processes that pool over a wide range of spatial frequencies and orientations (Morrone, Burr & Maffei,

1982; Bonds, 1989). Thomas and Olzak (1997) more fully describe tests of nonlinear gain control models that influenced the particular form of the model we present here.

2. Background

Our initial evidence for the involvement of higher-level processes in complex pattern discrimination tasks was demonstrated in a series of experiments designed to assess limitations of a class of psychophysical models developed for simple pattern detection tasks (Thomas, 1989; Thomas & Olzak, 1990; Olzak & Thomas, 1991, 1992). In this original class of model, which we termed *multiple channels*, *direct-access models* (Olzak & Thomas, 1992), it was assumed that observers base psychophysical decisions on information directly available from parallel, low-level analyzers at the initial level of cortical processing (V1). More formally, these models presume that the decision process has direct access to information represented by individual pathways or neurons in primary visual cortex.

Several decades of psychophysical and physiological research have determined that at the initial cortical level of processing, neurons jointly tuned to local spatial frequency and orientation information perform something akin to a patchwise Fourier analysis of a pattern. Physiological evidence suggests that different spatial frequencies and different orientations are locally represented in separate and independent neural pathways (Lennie, 1980; Robson, 1980; DeValois, Albrecht & Thorell, 1982; Enroth-Cugell & Robson, 1984; Shapley & Lennie, 1985; Van Essen, 1985; DeValois & DeValois, 1988). Behaviorally, an abundance of masking, adaptation, and subthreshold experiments have shown that such disparate stimulus components are detected independently (for reviews, see Olzak & Thomas, 1986; Graham, 1989). The multiple channels, direct-access models of detection provide an elegant link between visual performance and the underlying physiological mechanisms (Thomas, 1970; Wilson & Bergen, 1979; Watson, 1983), and it has been shown repeatedly that models of this class can be extended to account for discrimination performance among simple and some complex patterns (Thomas, 1970; Thomas & Shimamura, 1975; Thomas & Gille, 1979; Olzak & Thomas, 1981; Watson & Robson, 1981; Thomas, 1983; Watson, 1983; Wilson & Gelb, 1984; Klein & Levi, 1985; Nielsen, Watson & Ahumada, 1985; Thomas, 1985; Watt & Morgan, 1985; Wilson, 1986; Thomas & Olzak, 1990).

In more recent experiments, however, we reported gross failures of direct-access models to provide a general account of discrimination performance, even with rather simple patterns that contained two distinguishable sinusoidal gratings (Olzak & Thomas, 1991, 1992). Masking and cue-summation experiments revealed multiple (but not universal; see Thomas & Olzak, 1990) instances of strong interactions between grating components that differed widely in spatial frequency or orientation. Unexpectedly and importantly, some interactions observed were found to depend upon the dimension upon which the discrimination judgment was made (small differences in spatial frequency, orientation, or contrast) rather than upon the particular stimulus combinations. This finding implies that the neural mechanisms used in distinguishing one stimulus from another may be organized to signal specific types of information rather than simply responding to the presence of particular combinations of stimulus features.

The discrimination model we present in Section 3 builds upon earlier physiologically-tied, direct-access accounts of detection and discrimination. In developing the model, we attempted to provide a single theoretical framework within which we could account for the pattern of interactions and independence observed over the set of 12 complex-discrimination experiments we have previously reported (Thomas & Olzak, 1990; Olzak & Thomas, 1991, 1992), while maintaining compatibility with previous empirical results obtained by ourselves and others.

3. The model

The model has three coding/recoding stages and a decision stage, illustrated schematically in Fig. 1.

In Stage 1 of the model, the stimulus is initially analyzed by a bank of spatially tuned linear filters. In the second, nonlinear stage of processing, the output of each filter is transformed both by a within-pathway nonlinear process and by a divisive pooled gain control process. The gain control factor has been shown to be directly determined by the outputs of the linear filters in

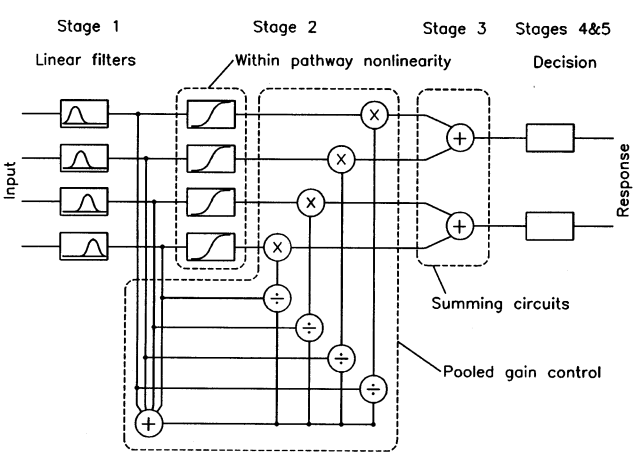

Fig. 1. Model overview. See text for description.

grating discrimination tasks similar to those described in the present paper (Thomas & Olzak, 1997). Thus, in Fig. 1, we show the output of the linear filter transformed by both within-pathway and pooled normalization processes. The processes operate in parallel and are then multiplied together to determine overall contrast gain. The normalized output of certain pathways are then selectively summed by higher-level summing circuits (only two are shown here) to signal certain types of information, such as the orientation of an edge or the textural grain of a pattern. Decision processes act directly upon the output of the appropriate summing circuits, which are chosen by the observer depending upon the nature of the information needed to perform the task at hand. In some instances, no summing circuit exists to combine responses to different stimulus components, and decision processes may take input transparently passed from normalized primary-layer units (not shown in Fig. 1).

We conceptualize the first three stages as a series of neural recoding processes that operate automatically and that provide information upon which perceptual decisions about complex stimuli can be made. The output of this three-part recoding module may be used to mediate a number of different perceptual tasks, which may require different models at the decision stage. The decision processes we describe here are specialized to link the output of the third processing stage to performance in a two-alternative pattern discrimination task. Eqs. (1)–(5) in Section 3.1 describe processing at each stage.

3.1. Neural recoding stages

Stage 1. In our notation[1], $R^{(x)}$ represents the output of neurons in stage x of the model. Thus, $R_i^{(1)}$ and $R_i^{(2)}$ represent activity within one neural pathway i, but at different levels or stages. The first layer of processing is an array of local neural pathways, each jointly tuned with respect to spatial location (x, y), spatial frequency, orientation, and phase. In Stage 1 of the model, we represent the output of each pathway i as the inner product of the sensitivity function of the filter $S_i(x, y)$ and the local luminance function $L(x, y)$:

$$R_i^{(1)} = \iint S_i(x, y)L(x, y)\mathrm{d}x\ \mathrm{d}y \quad (1)$$

Once defined in the space domain, each primary-level unit, i, is also localized in the Fourier domain. Thus, the subscript i signifies a single pathway. Different pathways vary in tuning characteristics with respect to spatial location, spatial frequency, orientation, phase, etc.

[1] The notation we adopt in this paper follows a system suggested to us by Thomas D. Wickens.

We assume a priori that the output of each unit is either zero or positive, in keeping with known cortical neurophysiology. In some recent models of pattern vision, this step is explicitly achieved via half-wave rectification (e.g., Solomon & Sperling, 1994; Sperling et al., 1994). In our model, we achieve the equivalent of half-wave rectification by assuming that the population of units included in the primary layer of processing is densely packed and varied enough with respect to phase to produce positive responses to both increments and decrements of light.

Stage 2. The second processing stage describes the nonlinear processes that control the overall response rate of each neural pathway. Our representation of the nonlinear stage is a modification of normalization formulations proposed by Legge and Foley (1980), Albrecht and Geisler (1991), and Heeger (1991, 1993). Our model combines two independent nonlinear processes, which have been described in slightly different notation in an earlier publication (Equation 2a, Thomas & Olzak, 1997). One process is a within-pathway nonlinearity that takes the form of a hyperbolic ratio with semisaturation constant c_1 and exponent p:

$$R_i^{(2')} = \frac{[R_i^{(1)}]^p}{{c_1}^p + [R_i^{(1)}]^p}$$

As Thomas and Olzak (1997) point out, the hyperbolic ratio both describes the way in which detection and discrimination performance vary with contrast (Thomas, 1983) and the contrast response function of many individual cortical neurons (Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982). A value of $p = 2.0$ approximates average results for the types of stimuli we have used in our studies.

The second process is a divisive gain control, or normalization process that independently acts to reduce the response of each neuron by a factor governed by the total activity in a pool of neurons:2

$$R_i^{(2'')} = \frac{[R_i^{(1)}]^r}{\left[{c_2}^q + \sum_j w_i(j)[R_j^{(1)}]^{rq}\right]^{1/q}}$$

The parameter c_2 is a constant that influences the contrast at which discrimination accuracy approaches an asymptote. The pool over which the normalization process occurs is defined by the weighting function $w_i(j)$, centered on the location of unit i, and operating over a set of units j, of which i is a member. Although we presume that the weighting function falls off gradually as a function of the similarity between tuning characteristics of unit i and other primary-layer units, we do not explicitly assume a form for the function $w_i(j)$. Instead, the function is empirically determined in masking experiments that isolate this stage of processing. The weighted responses of the primary-level units included in the gain control pool are combined

using a Minkowski-type metric. The parameter q indicates how responses from units are combined within the gain control pool.

We include the parameter r for generality, although a value of 1.0 has been found to be consistent with empirical results for the type of tasks we describe here (Thomas & Olzak, 1997). When $r = 1$, the pooled gain control processes operate directly on the output of the linear filter, as we have depicted in Fig. 1. When $r = 2$, the equation describes a contrast energy model, and operates upon the squared output of filter responses (Heeger, 1991, 1993).

The pool we describe in the normalization process explicitly includes neurons tuned to a wide range of frequency and orientations, based on evidence from our earlier experiments (Olzak & Thomas, 1991, 1992). In our discrimination data, a normalization process that operates over a broad, but ultimately limited pool of primary-layer neurons was found to be necessary to account for the cross-orientation and cross-frequency masking we observed (Olzak & Thomas, 1991, 1992). As suggested by Debruyn and Bonds (1986), one possible physiological substrate for such a process is intracortical inhibition, which has been reported to occur over a wide range of orientations (Morrone, Burr & Maffei, 1982; Bonds, 1989) and spatial frequency (Bonds, 1989).

Pooled normalization processes have been proposed in the context of a number of recent models of texture segregation, motion, and other visual processes (Grossberg & Mingolla, 1985a,b; Bergen & Landy, 1991; Graham, Beck & Sutter, 1992; Wilson & Humanski, 1993; also see Graham, Beck & Sutter, 1992; Wilson, 1993; Graham & Sutter, 1998 for discussions of different forms of nonlinearities), but other formulations have focused on pooling in the spatial domain; i.e. pools that include neurons with similar Fourier tuning characteristics but different tuning with respect to spatial location. Our formulation permits pooling in the spatial dimension as well. However, because we have thus far only investigated patterns with spatially overlapping components, spatial pooling characteristics of the processes revealed in our discrimination tasks have not yet been empirically determined. We therefore do not discuss spatial pooling further in this paper.

Because gain control alters the height of the log performance versus log contrast function but not its shape (Thomas & Olzak, 1997), we assume that the within-pathway nonlinearities and the gain control nonlinearities combine multiplicatively. The output of the combined second stage is:

$$R_i^{(2)} = \frac{[R_i^{(1)}]^r}{\left[c_2{}^q + \sum_j w_j(j)[R_j^{(1)}]^{rq}\right]^{1/q}} \times \frac{[R_i^{(1)}]^p}{c_1{}^p + [R_i^{(1)}]^p} \quad (2)$$

Stage 3. The specialized families of summing mechanisms we have isolated in our discrimination studies are characterized at the third stage of the model:

$$R_\alpha^{(3)} = \left(\sum_i v_\alpha(i)[R_i^{(2)}]^{q'}\right)^{1/q'} \quad (3)$$

In Eq. (3), normalized responses of units, $R_i^{(2)}$are combined in parallel by multiple third-stage summing circuits. For each third-stage unit α, the output $R_\alpha^{(3)}$ is determined by a characteristic weighting function $v_\alpha(i)$. The weighting function describes the particular neural pathways i that are included in the summation and defines the bandwidth characteristics of the third-stage mechanism with respect to the dimensions of spatial frequency, orientation, and phase[2]. Reflecting the interpretation of our earlier findings, which we describe in the remainder of this paper, the subscript α in Eq. (3) indicates that the weighting function, or bandwidth profile, depends upon the type of information represented.

The value of the exponent q' describes how responses of units are combined within a third-stage mechanism[3]. We have denoted the summing exponent q' in Eq. (3) to reflect an empirical result that requires its value to be identical to the summing exponent q in Eq. (2), although conceptually the two summations are different and in principle could take on different values (Thomas & Olzak, 1997). The constraint is based on the result that when measured as a function of contrast, discrimination performance reaches the same asymptote regardless of whether the stimuli each contain a single component that acts as a cue or two components that both act as cues. If the summing exponent associated with the masking pool in Stage 2 was lower in value than that associated with the summing circuits in stage 3, then performance in the two component case would asymptote at a lower level than that in the single level case, and vice versa. An analogous argument holds for how performance varies with stimulus size.

As shown in Appendix A, the experiments discussed in this paper are insensitive to the absolute values of q and q'. To simplify our exposition, we assume values of 1.0 for the purposes of this paper.

3.1.1. Compatibility with previous models

When the two stimuli to be discriminated each comprise only a single component, pooled gain control and higher-level summing processes become transparent. Earlier direct-access models assumed a simple linear

[2] Additional dimensions, such as spatial summing characteristics or tuning with respect to color, can also be specified.

[3] We use a Minkowski combining function to describe the combining process. See Graham (1989) for a discussion of the Minkowski metric and interpretations of combining exponent values.

filtering stage followed by a within-pathway nonlinearity (Thomas, 1970; Wilson & Bergen, 1979; Watson, 1983). The within-pathway nonlinearity is necessary to account for the function relating simple discrimination performance to stimulus contrast for single-component stimuli (Thomas, 1983). In the model described here, we have included an additional pooled gain control process in the second stage of the model, as well as a higher level summing stage which may be linear or nonlinear. When only a single stimulus component is present, the model reduces to the algebraic form of earlier models. Our model is therefore entirely compatible with previous direct-access formulations that accounted for simple discrimination data.

3.2. Decision stages

Stage 4. We link our recoding model to psychophysical performance in the fourth and fifth stages of the model. We assume that observers do not have direct access to the primary processing layer when making discrimination decisions, but use the output of an appropriate third-stage mechanism to distinguish among stimuli. If no mechanism exists to combine two particular components of a stimulus, decision processes may combine transparently-passed information from two or more third-stage mechanisms.

We conceptually separate the decision stage of the model into a separate module to emphasize the possibility that responses from third-stage mechanisms may play roles in visual tasks other than the discrimination judgments discussed in this paper. Such possibilities can be investigated by substituting appropriate decision rules at the decision stages.

Discrimination judgments of the sort described here are made by differencing the outputs of the two summing circuits, $R_A^{(3)}$ and $R_B^{(3)}$, with (1) the greatest differential sensitivity to two patterns to be distinguished, A and B, and (2) that belong to the family that is specialized to provide the type of information upon which the discrimination is based. We note that although each summing mechanism alone may be differentially sensitive to the two patterns, the output of a single mechanism is ambiguous if contrast of the patterns vary. We overcome this ambiguity by using a differencing scheme between two mechanisms to generate a single decision variable $R^{(4)}$:

$$R^{(4)} = R_A^{(3)} - R_B^{(3)} \quad (4)$$

Although the differencing operation is not necessarily the optimal decision rule for discrimination, this simple scheme has been successfully used both in past (Thomas & Shimamura, 1975) and current (Graham, Beck & Sutter, 1992) models of spatial vision, and appears to adequately describe performance in discrimination tasks (see Graham, 1989 for a full discussion of decision rules).

Stage 5. In the final stage of the model, we link the decision variable $R^{(4)}$ to the signal-detection performance measure d', which we use in our empirical work. This measure is the difference between the fourth-stage response to stimulus A and the fourth-stage response to stimulus B, divided by the S.D. of the $R^{(4)}$ responses:

$$d' = \frac{R(A)^{(4)} - R(B)^{(4)}}{\sigma} \quad (5)$$

We recognize that processing at each stage up to this point is perturbed by random variability which may or may not be normally distributed. However, ROC curves that we have obtained over the years suggest that the underlying decision variable has a Gaussian distribution. Consistent with this observation, we represent noise only at the final decision stage, and assume that the $R^{(4)}$ response is approximately normally distributed with unit variance independent of which stimulus was presented (equal-variance assumption). Smith and Thomas (1989) found the equal-variance assumption to be valid for discrimination tasks of the type described here.

4. Linking model and tasks

In Sections 4.1, we briefly review the discrimination tasks used in our previous experiments and the predictions of direct-access independence models. In Section 4.2, we schematically illustrate the theoretical logic of our earlier experiments to clarify how any interaction must arise either from decision processes themselves or from processes interposed between the primary level and the decision process.

4.1. Stimuli and procedures: testing the direct access model

The set of experiments and data summarized in this section were originally reported in Thomas and Olzak (1990), and Olzak and Thomas (1991, 1992). These studies tested the adequacy of the direct access models, which presume only a primary layer of processing and a decision process that has direct access to responses from the primary layer. The data were gathered in studies in which the viewer always discriminated between two highly similar patterns, which differed slightly from one another in either spatial frequency, orientation, or contrast. The stimuli in all experiments were patches of suprathreshold sinusoidal gratings ($10 \times$ individual contrast thresholds), windowed spatially by a bivariate Gaussian function with a space constant of 40 min (Gabors). Gabor stimuli were chosen because they are localized in both the spatial and the spatial frequency domains; thus, each windowed grating stimulates only a relatively small number of

neurons at the primary level of processing (DeValois & DeValois, 1988). Each experiment was designed to test whether, in suprathreshold discrimination tasks, two superimposed gratings of very different spatial frequency and/or orientation were processed independently. The components were chosen to activate separate and independent groups of neurons at the primary layer of processing (DeValois & DeValois, 1988; Graham, 1989). Different experiments tested whether conclusions about independence depended (1) upon the aspect of the stimulus upon which the discrimination was based (contrast, spatial frequency, or orientation) and (2) upon whether components differed greatly in spatial frequency, orientation, or both.

Before introducing the different experiments and conditions, it is helpful to understand the discrimination paradigm that was used throughout the series. In each block of trials, each of two patterns to be discriminated, A and B, was presented 50 times, randomly intermingled in a block of 100 trials. In control conditions, each pattern comprised a simple Gabor stimulus. In masking and cue-summation conditions, each stimulus contained two components. Although the two complex patterns to be discriminated in masking and cue-summation conditions were always very similar, the components comprising each pattern differed widely in spatial frequency, orientation, or along both dimensions. On a given trial in any condition, the viewer saw only one of the two possible patterns on a given trial and rated his or her certainty that pattern A or pattern B had been presented. The judgment was made using a 6-point rating scale, in which a '1' indicated certainty that stimulus A had been presented on a given trial and a '6' indicated certainty that stimulus B had been presented. Intermediate values indicated varying degrees of certainty.

The measure of discrimination accuracy, d', was calculated by one of two methods. For some experiments, we constructed a rating ROC from the responses and the area under the ROC was estimated by the polygon method. The area was transformed to a standard-normal deviate (z-score) and multiplied by $2^{0.5}$ to estimate d' (Thomas & Olzak, 1990; Olzak & Thomas, 1991). In other experiments, d' was calculated by fitting a normalized ROC of unit slope to the rating data, a procedure that minimizes error on both axes, assuming equal variance (Olzak & Thomas, 1992). For each observer, the d' values from five to eight daily sessions were used as input to an analysis of variance in order to make statistical comparisons across conditions.

Within each experiment, six conditions were run, each in a separate block of trials: two control conditions, two masking conditions, and two cue-summation conditions. Fig. 2 shows examples of control, masking, and cue-summation conditions in the top, middle, and bottom rows, respectively. In control conditions, each of the two stimuli to be discriminated contained a single grating component that differed slightly from the other in orientation, spatial frequency, or contrast, depending upon the experiment. The top row of Fig. 2 shows an example of two pairs of stimuli used in the two control conditions of one experiment, designed to test for independence over different frequency bands when the task was a discrimination based on small differences in orientation[4]. For clarity, the size of the differences shown in Fig. 2 are approximately $10 \times$ actual differences used in running the experiments. The left-hand pair (low frequency control condition) consists of 3 cpd gratings that differ slightly from one another in orientation. Pattern A tilts slightly to the left of vertical; pattern B tilts slightly to the right. The right-hand pair also differ from one another in orientation, but the gratings are centered on a much higher frequency, 15 cpd (high frequency control condition). The grating pairs were always chosen to be disparate enough to stimulate separate and independent sets of units at the primary layer of processing (Olzak & Thomas, 1986; Graham, 1989 for reviews of the evidence). In the example shown in Fig. 2, the components are drawn from two different frequency bands. In other experiments, the components were similar in frequency but superimposed at orthogonal orientations. In a third type of experiment, the component could differ in both spatial frequency or orientation.

In masking and cue-summation conditions, each of the two stimuli to be discriminated contained two superimposed gratings, one from each frequency (and/or orientation) band. Masking experiments asked how the presence of the second, constant component affected the ability to discriminate between the two patterns relative to controls. Cue-summation experiments asked how information represented in two very different frequency and/or orientation bands was combined. Both masking and cue-summation studies tested whether information presented by the two superimposed gratings was processed independently.

In our example, the stimuli shown in the middle and lower panels of Fig. 2 each contain two component gratings, one at 3 cpd, the other at 15 cpd. In the two masking conditions, (middle row, Fig. 2), each control discrimination was repeated, now in the presence of a constant mask drawn from the 'other' frequency band. The mask always took an intermediate value along the dimension to be discriminated (e.g. vertical in this example). Thus, in the left-hand pair of gratings, the 3 cpd grating still tilts left in pattern A and right in pattern B by the same amount as it did in controls, but each grating is now summed with a vertical, 15 cpd

[4] Examples of stimuli used to test for independence across orthogonal orientations can be found in Figure 1 of Olzak and Thomas (1991).

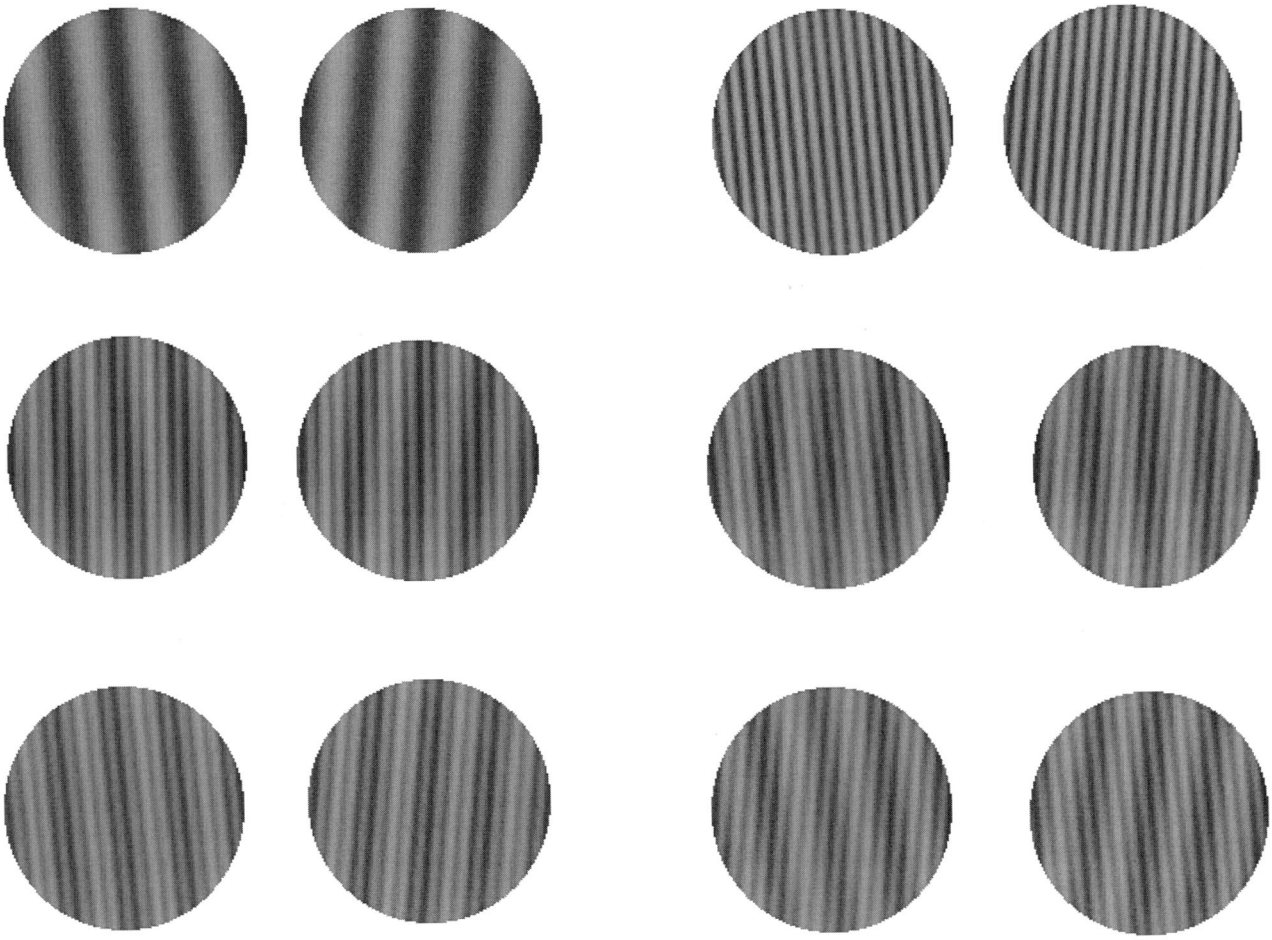

Fig. 2. Examples of stimulus pairs used in control (top row), masking (second row), and cue-summation (third row) conditions in an experiment investigating how orientation information is combined over widely-separated frequency bands Note that actual stimuli were spatially windowed by a Gaussian, and differences with respect to orientation were much smaller in actual running conditions.

mask. In the right-hand pair of gratings, the 15 cpd gratings vary in orientation as they did in controls, but each is summed with a vertical 3 cpd mask. The components are processed at the primary level of processing via separate and independent neural pathways. If decision processes have direct access to these pathways, the mask should have no effect on discrimination performance relative to control.

The cue-summation test we use was adapted from the phase test used to assess independence of processing in several domains (Boynton, Ikeda & Stiles, 1964; Graham & Nachmias, 1971; Cohn & Lasley, 1976). In cue-summation conditions, both the low and high frequency components of each stimulus to be discriminated provided a cue to discrimination. Two cue-summation conditions were tested that differed in how the cues were combined. The bottom row of Fig. 2 illustrates the two cue-summation conditions for the example at hand. In both the left-hand and right-hand pairs of stimuli, the 3 cpd grating tilts to the left in pattern A and to the right in pattern B. The 15 cpd grating also tilts (by the same amount) in each pair, providing a second cue upon which to base the discrimination. However, the pairs differ with respect to the direction of the tilt of the second component. In the left-hand pair, the 15 cpd grating tilts the same way as the 3 cpd grating (left in pattern A, right in pattern B). This is termed the *same-sign* condition. In the right-hand pair, the 15 cpd component tilts in opposition to the 3 cpd component. We term this the *opposite-sign* condition[5]. In both conditions, each component differs by the same amount between the two stimuli to be discriminated, regardless of how they are combined. Again, the 3 and 15 cpd components are processed independently at the primary layer of processing. If decision processes have direct access to these pathways, performance in the two conditions will be equal. We term any difference in performance between the two cue-summation conditions a *configural effect*, because performance depends upon whether the two cues to

[5] In some earlier publications, the same-sign and opposite-sign conditions were termed positively and negatively correlated conditions, respectively.

discrimination are combined in the stimulus in same- or opposite-sign fashion.

As noted earlier, component pairs differed in different experiments, as did the judgment dimension. We ran a total of 12 experiments, which can be seen in summary form by looking ahead to Fig. 5. Each experiment is represented jointly by the stimulus type (rows) and judgment dimension (major columns). The experiments listed in the top row tested for independence of processing when components differed in both spatial frequency and orientation (e.g. a 15 cpd horizontal grating superimposed upon a 3 cpd vertical grating). Experiments listed in rows 2 and 3 tested for independence across different orientations when components were of similar spatial frequency. Stimuli were formed by superimposing orthogonal gratings, both either 3 cpd (second row) or 15 cpd (third row). Experiments listed in row 4 tested for independence processing over different spatial scales (e.g. the $3+15$ vertical gratings as shown in our example).

For each of these three types of stimuli, assessments were made for three different types of stimulus information by discriminating of the basis of small differences in spatial frequency, contrast, or orientation (columns 1, 2, and 3, respectively).

4.2. Theoretical logic of the experiments

To facilitate an understanding of the logic of the experiments, and to later represent characteristics of the summing mechanisms isolated in our experiments, we introduce here a schematic representation of the primary processing layer of our model, and show how our experiments tested for interactions across components that differed greatly from one another. A more mathematical treatment of predictions based on a multiple-channels, direct-access model can be found in Olzak and Thomas, 1992. Here, we simply develop the logic schematically.

Fig. 3 shows an array of the linear mechanisms we assume in the first stage of the model, plotted in a polar-coordinate representation of the spatial frequency-orientation plane in the Fourier domain. Spatial frequency is represented as a distance from the origin; orientation is represented by the sweep angle. Small regions of this space represent response characteristics of local linear units similar to neurons found in V1, shown by the orientation and frequency at each small region. All are tuned to the same patch in the visual field.

Any small region in the plane can represent a single-component Gabor stimulus, and can also represent the response characteristics of local units that respond to it. The original null hypothesis to be tested in our experiments was that fine spatial discriminations are mediated by independent sets of tuned neurons at the primary level of processing. To continue our example, consider first the control patterns illustrated in the top row of Fig. 2. Each presentation of control patterns A or B excites a limited number of local linear units at the primary level of processing. Units responding to the two patterns are shown in the overlapping circles labeled A and B in Fig. 3. Both sets of units respond to nearly vertical, low-frequency patterns. Because the difference between patterns A and B is small, A and B excite nearly identical groups of cells with highly overlapping sensitivities at the primary level of processing.

Now consider adding a second, very different component to each of the stimuli to be discriminated. The second component, regardless if it acts as a mask or provides a second cue to discrimination, also activates only a small set of neurons, which are represented in a very different location in the spatial frequency-orientation plane. It is widely accepted that at the primary layer of processing, the components we chose to use in our tests (3 and 15 cpd, and/or orthogonal orientations) each activate a separate and independent set of primary-layer units (DeValois & DeValois, 1988; Graham, 1989, for reviews of the evidence).

In Fig. 3, we have illustrated the location of responses to the masking component shown in the second row of Fig. 2 by the circle labelled M. Notice that the mask is represented as exactly vertical, halfway between the values of the two lower-frequency components to be discriminated. The figure illustrates that the population of units responding to the mask is nonoverlapping with the units responding to the control components with

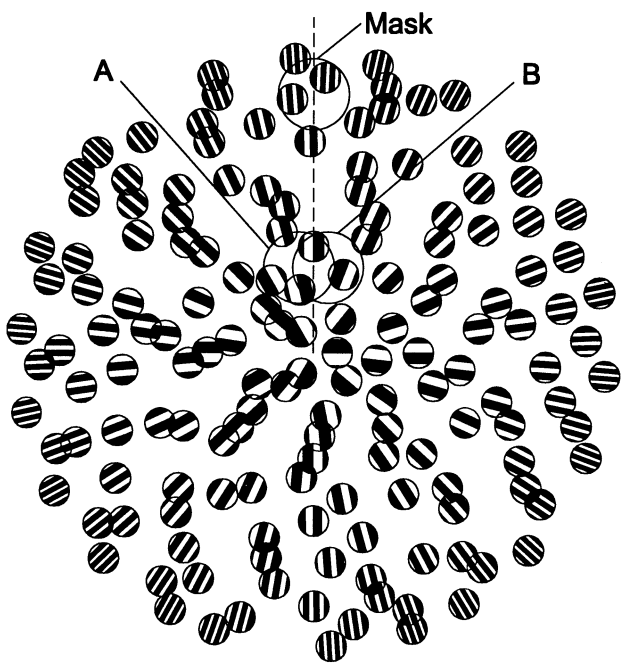

Fig. 3. Schematic diagram of primary layer units in a polar spatial frequency-orientation (adapted from DeValois, Albrecht & Thorell, 1982). See text for description.

respect to spatial frequency. Thus, any interactions must occur beyond the primary layer of processing, either at the decision stage of processing or through processes interposed between the primary layer of processing and the decision stage. A similar logic applies to the cue-summation studies when the second cue added to each stimulus is drawn from a very different frequency and/or orientation band.

4.3. Results of previous experiments

The results of our 12 earlier experiments took one of two fundamental forms. In some instances, independence obtained. Neither masking nor configural effects were observed, and performance was approximately equal in all conditions. In other experiments, a characteristic pattern of nonindependence was revealed in both masking and cue summation conditions. Whether independence or interaction obtained depended both upon the type of judgment being made (spatial frequency, orientation, or contrast) and whether the stimulus components differed widely in frequency or orientation. We discuss the pattern of independence/interaction over experiments in Section 5. Here, we focus on the two typical patterns of results within an experiment.

Fig. 4 illustrates the two typical patterns of results that we observed. A key point is that both sets of data were collected with stimuli virtually identical to those shown in Fig. 2. The two components comprising the complex stimuli in each experiment were at or near vertical. One component was at or near 3 cpd; the other was at or near 15 cpd. Thus, both of these experiments tested for independence of processing over different frequency bands. They differed only in whether the discrimination judgment was made on the basis of small changes in spatial frequency or in orientation.

Panel A shows results when the task was a spatial frequency judgment, averaged over three observers. Conditions C_3 and C_{15} show performance obtained with the 3 cpd and 15 cpd controls. Conditions M_3 and M_{15} show performance when a 15 cpd or a 3 cpd mask was added to the 3 cpd and 15 cpd gratings, respectively. Conditions SS and OS show performance in two-cue conditions when components were added in same-sign and opposite sign configurations. Performance in all conditions was approximately equal, with no significant differences found for any observer. This pattern of results is consistent with predictions of direct-access models, although the data suggested an inability to simultaneously use information from both frequency bands.

Panel B shows results for the nearly identical stimuli when the task was an orientation judgment, and illustrates the classic pattern we obtained whenever we found evidence for interactions. Relative to controls,

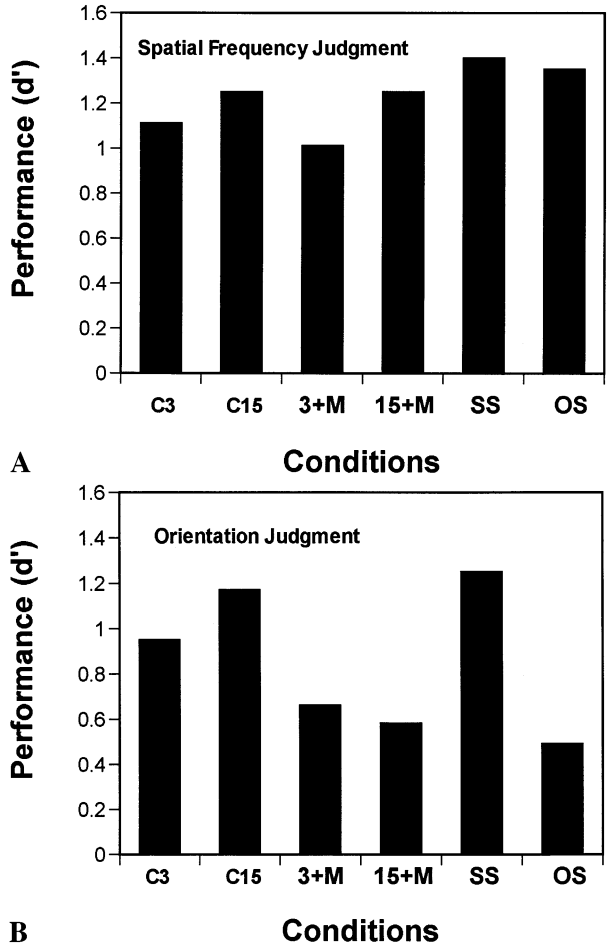

Fig. 4. Example results showing the two typical pattern of results, both obtained with 3 + 15 vertical components. Panel A: no interactions across frequency bands when discrimination is based on spatial frequency (data from Thomas & Olzak, 1990). Panel B: masking and configural effects reject direct-access models when discrimination is based on orientation (data from Olzak & Thomas, 1992).

performance dropped to nearly half by the addition of a mask that differed in frequency by a factor of five. The masking effect disappeared when the second component provides a cue to discrimination, but only when the cues rotated together (SS vs OS). In this case, both the masking effect and the presence of a configural effect (i.e. finding that how cues were combined affected performance levels) strongly rejected direct access models.

In the 12 previous experiments discussed here, whenever we observed a configural effect, we also found masking effects, although in one instance (orientation judgments on similar-frequency plaids) the magnitude of masking was very small[6]. In cases showing

[6] In other experiments we have performed to measure bandwidths of mechanisms underlying masking and cue-summation effects, we have found masking in cases when no configural effects were observed (Olzak & Thomas, 1995, 1996).

interactions across disparate frequencies or orientations, the presence of a mask reduced performance to approximately one-half that in control conditions. In cue-summation conditions, observed configural effects always yielded better performance when the second cue corresponded to the first in sign (same-sign condition; e.g. both tilted left). In same-sign conditions, performance was approximately twice that found in masked conditions, a point we return to in Section 6. In opposite-sign conditions, (e.g. one component tilted left and the other right), performance was as poor or even worse than under single-cue masked conditions.

The classic (and recurrent over experiments) pattern of interaction we observed provided the basis for developing the general structure of the model presented in Section 3, and for making certain inferences about the nature of processing at each stage.

Masking by components that stimulate no overlapping pathways at the primary layer of processing cannot occur in a system that only includes within-pathway nonlinearities. Thus, there must be some other process that modifies the response of a given pathway by activity in pathways with very different tuning. There are a number of biologically plausible vehicles to accomplish this, including signal-dependent noise (Thomas, 1983; Olzak & Thomas, 1992; Geisler & Albrecht, 1995) or intracortical inhibition (Bonds, 1989; Morrone et al., 1982). We have chosen to model the masking with an adaption of well-specified formulations of pooled gain control processes already in the literature (Albrecht & Geisler, 1991; Heeger, 1991, 1993), and have shown elsewhere (Thomas & Olzak, 1997) that the formulation we propose, in which pooled gain control processes act upon the direct output of the linear filter, is quantitatively consistent with data indicating how discrimination performance varies with contrast and size.

Within the context of the model, effects due to gain control are found by comparing discrimination performance in control conditions (Fig. 2, top row) to that obtained in masking conditions (Fig. 2, middle row). In control conditions, the response of each pathway is modified only by activity in the small group of primary-layer units with similar tuning characteristics that respond to the simple control grating. The addition of a mask that differs greatly from the control grating in terms of spatial frequency or orientation produces activity in a different set of primary-level units. If the units responding to the mask are not a part of the gain control pool described in the denominator of Eq. (2), then performance should be unaffected by the addition of the mask. However, if the units responding to the mask are a part of the gain control pool, then they contribute to the normalizing process and the response of each unit R_i is reduced relative to its response in control conditions. This, in turn, reduces the difference in Eq. (4), resulting in decreased performance in masking conditions relative to control[7].

In the context of the model, the effects of third stage summing circuits are isolated in the cue-summation conditions (Fig. 2, bottom row) and in comparisons among performance in the two cue-summation conditions and in masking conditions. In masking conditions (Fig. 2, middle row), only a single cue to discrimination is present; the second component adds no differential information to the stimuli to be discriminated. In cue-summation conditions, each stimulus again comprises two components, but now both components provide potential cues to discrimination. An important feature of the masking and cue-summation comparisons is that the effect of gain control is held essentially constant from one stimulus to another, because both patterns contain two components[8]. Consequently, differences in performance among cue-summation and masking conditions reflect only the operations of third stage summing mechanisms.

The relationships observed between performances in the masking and cue-summation conditions lead to inferences about the nature of the information available at the decision stage when making fine spatial discriminations. When performance in the two conditions is identical, we cannot reject the direct access model, and decision processes treat each component as an independent source of information. Any increase in performance over masked levels indicates either probability summation or some type of information integration. For example, if the observer integrates the information from these independent sources according to an ideal-observer model, performance in both conditions will be the Euclidean sum of the two d' values obtained in masking conditions (Olzak & Thomas, 1991, 1992).

Unequal performance between the same-sign and opposite-sign conditions indicates that observers are

[7] If the third stage of processing is not linear, then this comparison jointly reflects processing at the second and third stage of processing. However, we have determined in other work that bandwidths of gain control processes are quite broad with respect to both frequency and orientation, whereas bandwidths of third-stage summing circuits are quite narrow along the dimension they signal (Olzak & Thomas, 1995, 1996; Thomas & Olzak, 1997). Thus, the two stages of processing are experimentally separable, leading us to propose both stages.

[8] Strictly speaking, the gain control terms should be slightly higher in the same-sign condition, where both components of each stimulus have exactly the same orientation, than in either the masking or in the opposite-sign condition, where the two components of each stimulus differ slightly from each other in orientation. However, as has been shown elsewhere (Thomas, 1989; Olzak & Thomas, 1992; Thomas & Olzak, 1997) the gain control weighting function changes relatively slowly as a function of spatial frequency and orientation. Consequently, the differences between corresponding gain control terms are so small as to be safely ignored. Because gain control is held constant, any differences in performance among the masking and cue summation conditions reflect activity only at the third stage of processing.

basing their decisions on information that has been combined across the two components comprising each stimulus to be discriminated prior to the decision stage of the model. Because the components are chosen to activate separate and independent units at the primary stage of processing and the effects of gain control are held constant, the combining process can only occur at a different processing stage than gain control.

The presence of a configural effect indicates that the combining process takes into account the sign (e.g. direction of tilt) of the cue in each stimulus. The direction of the configural effect provides information about the nature of the combining process that is occurring at the third processing stage. If the decision is based on using one component as a reference against the other, then performance will be best in the opposite-sign condition. If the decision is based upon a summation (effectively signalling an average value) of responses to the two components, then performance will be best in the same-sign condition. Our data indicated that when combining processes were observed, performance was always best in the same-sign condition. We therefore infer that the combining processes we have isolated in our cue-summation experiments are summing, rather than differencing, circuits.

5. Interpretation of experimental series

The example used to introduce our methods (Fig. 2) used stimuli that tested for independent processing of grating components that differed widely in spatial frequency when the task was an orientation discrimination. However, our model was developed to account for a series of such experiments. The full set of experiments examined how results depended upon the dimension along which the superimposed components differed (differences in spatial frequency band, orientation band, or in both orientation or frequency band). An overview of the general plan of the experiments can be found by examining Fig. 5.

The major columns of Fig. 5 show the three dimensions of judgment upon which fine discriminations were made (small differences in spatial frequency, orientation, or contrast). The rows indicate stimulus combinations. Three classes of stimuli were examined. In the first class, components differed in both spatial frequency and orientation, as shown in row 1. These experiments tested for interactions when the superimposed gratings differed in both the dimensions of frequency and orientation. The second class is shown in rows 2 and 3. These stimuli comprised two superimposed gratings that were horizontal and vertical and of the same frequency, either 3 cpd (row 2) or 15 cpd (row 3) and tested for interactions across orthogonal orientations when components were similar in spatial frequency. The third class of stimuli, shown in row 4, were formed by superimposing 3 and 15 cpd near-vertical gratings as in our example shown in Fig. 2. These experiments, as we have noted earlier, tested for interactions across different frequency bands.

5.1. Capturing previous results

Fig. 5 also summarizes our previously-reported empirical results with respect to masking and configural effects (Thomas & Olzak, 1990; Olzak & Thomas, 1991, 1992). In this section of the paper, we describe how, in the context of the model, we interpret these results. Where appropriate, we also summarize results of subsequent experiments that have clarified our interpretation of the data.

5.1.1. Second-stage processing: gain control

Consider first the pattern of masking effects across stimuli and judgment types in Fig. 5. We initially observed that when components differed widely in both spatial frequency and orientation (top row), no masking occurred in any task. These results indicated that there is no general, pooled normalization process affecting the whole of the primary processing layer. However, when components differed widely in orientation, but were of similar spatial frequency (rows 2 and 3), masking effects reliably occurred (although we note that masking effects were quite small when the judgment was made on the basis of small differences in orientation, columns 5 and 6). This result indicated the need to include a pooled gain control term as part of the nonlinear portion of the model, as discussed in Section 4. In particular, these two sets of results suggest

	Judgment Dimension					
	Spatial Frequency		Contrast		Orientation	
Stimulus	Masking	Configural Effect	Masking	Configural Effect	Masking	Configural Effect
3+15 Plaid	No	No	No	No	No	No
3+3 Plaid	Yes	Yes	Yes	Yes	Small but reliable	Yes
15+15 Plaid	Yes	Yes	Yes	Yes	Small but reliable	Yes
3+15 Vertical	No	No	No	No	Yes	Yes

Fig. 5. Pattern of results observed in experimental series. Cells indicate whether or not a masking or configural effect occurred in each of 12 different experiments. The dimension of judgment in each experiment is shown in major columns. Rows show the experimental stimuli used in each experiment plaids (rows 1–3) contained a superimposed horizontal and vertical component, one at each the indicated spatial frequencies. In row 4, both components were vertical, as shown in Fig. 2. Minor columns indicate the two types of effect possible in each experiment.

a normalizing pool that is broadband with respect to orientation, but limited with respect to spatial frequency.

Thomas and Olzak (1997) analyzed masking data gathered from several sources, including some of studies represented in Fig. 5, to estimate the orientation and spatial frequency bandwidths of the gain control pool. All data analyzed were gathered at high contrasts, where discrimination approaches an asymptote. At these contrasts, masking effects appear as reductions in asymptotic performance and reflect the effects of gain control. Masking effects declined as target and mask increasingly differed in spatial frequency, but no decline was found until test and mask components differed by more than 1 octave. Their results also suggested little tuning of the normalizing pool with respect to orientation, although orthogonal masks produced somewhat less masking than parallel masks. In terms of the model, these results suggest that the weighting function $w_i(j)$ in Eq. (2) declines slowly with differences in spatial frequency, and varies somewhat with orientation.

The interpretation of a single gain control pool with limited frequency bandwidth is supported by two of the three results shown in the bottom row of Fig. 5; when spatial frequency or contrast is judged on similarly oriented components of very different frequency, no masking is found across disparate frequencies. However, frequency-limited single-pool interpretation cannot account for the masking found with orientation judgments on $3+15$ vertical components (Fig. 5, row 4). The latter result may imply the existence of several normalization pools. Therefore, a second, though less parsimonious, interpretation of the pattern of masking results shown in Fig. 5 is that at least some of the normalization process is associated with task-specific third-stage mechanisms.

5.1.2. Third-stage processing: summing circuits

The interpretation of our data series with respect to third-stage summing units is straightforward, and has been supported in subsequent experiments (Thomas & Olzak, 1996; Olzak & Wickens, 1997). For ease of exposition, we first present schematics of the third-stage units, depicting summing profiles (with respect to spatial frequency and orientation) of the two complementary mechanisms that we inferred from our original data. We then develop the logic underlying our inferences and the need for multiple summing mechanisms specialized to signal different types of information.

In Fig. 6(A,B), the first-stage array is overlaid by schematics of the third-stage summing circuits revealed in the original data by our configuration-effect test. Panel A illustrates profiles of the mechanism type revealed when making orientation judgments. This family of mechanisms sums over all spatial frequencies within a limited orientation range, appearing as a fan of 'cigars' in this space. Mechanisms such as this are optimal to signal the orientation of real-world stimuli such as edges, lines, or oriented textures, which are simple stimuli in the space domain but complex in the Fourier domain. Because they sum over all frequencies, the mechanisms respond regardless of the exact frequency content, and therefore can signal the orientation

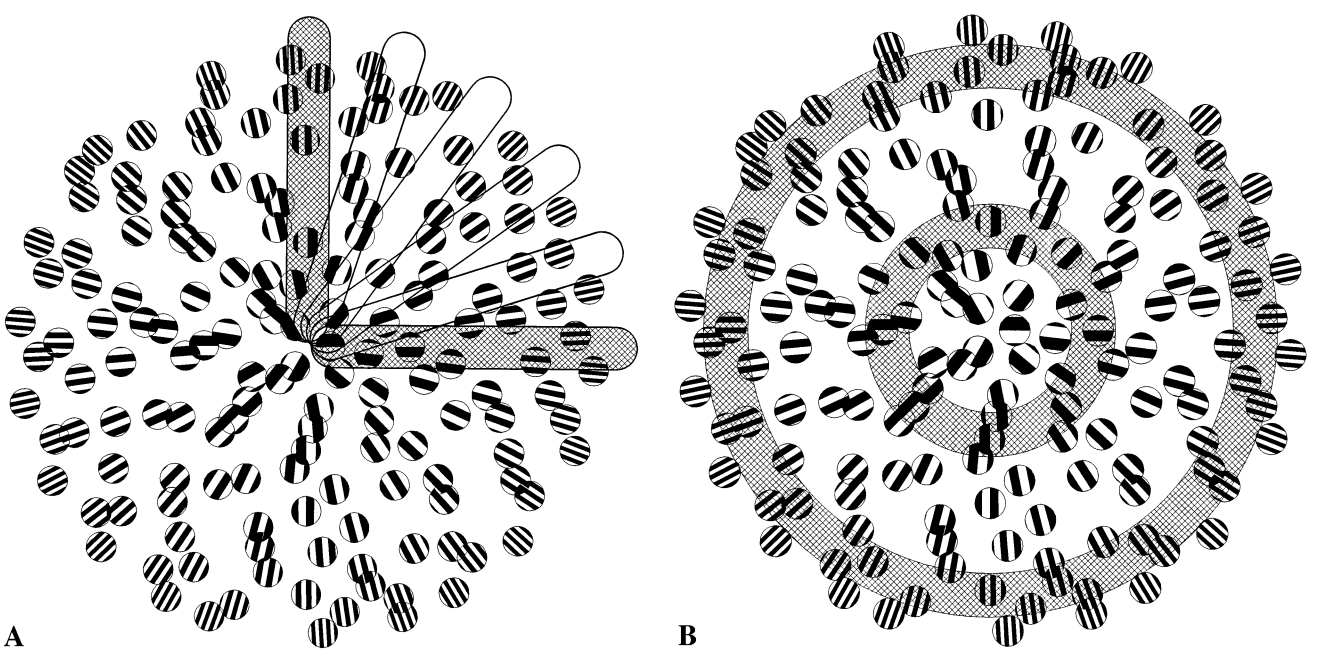

Fig. 6. Families of third order summing circuits specialized to signal orientation (panel A) and spatial grain (panel B). A mechanism similar to that shown in panel B sums contrast information.

of any stimulus feature regardless of its frequency content when it is presented at any arbitrary distance. In other words, they are invariant with respect to distance or retinal image size.

Figure Fig. 6(B) shows schematic profiles of the mechanism family (or families, as we shall argue below) revealed when making either spatial frequency or contrast judgments. Mechanisms signalling information about these characteristics sum responses over all orientations, within a limited frequency band. These mechanisms appear as concentric rings or 'doughnuts' in the spatial frequency-orientation plane, and appear suited to signal information about the spatial grain or contrast of either a one- or two-dimensional texture pattern. Because these mechanisms sum over orientations, they can signal information about the surface of a pattern regardless of component orientations. This invariance with respect to orientation makes these mechanisms ideal to signal information about the textural grain (or contrast) of surfaces viewed at any arbitrary orientation.

We now consider how the data lead to the third-stage mechanism profiles of Fig. 6. In the context of the model, the presence of any configural effect signals the existence of a third-stage summing circuit. We observed that when components differed widely in both spatial frequency and orientation, no configural effects occurred, regardless of the judgment dimension (Fig. 5, top row). Within the context of the model, this result suggests that there are no third-stage summing mechanisms that draw from units tuned to very different values of both spatial frequency and orientation.

When components differed widely along one dimension only (Fig. 5, rows 2–4), either orientation or spatial frequency, independence was found in some cases, configural effects in others. The result depended both upon whether the components differed in orientation or in spatial frequency and, in the case of $3+15$ vertical components, upon the dimension of judgment.

First, we observed that effects obtained with spatial frequency and contrast judgments (Fig. 5, columns 1–4) followed an identical pattern of results across experiments, but orientation judgments showed a different pattern (columns 5 and 6). This suggested the existence of two different summing circuits, each specialized to signal a different type of information about a stimulus. Consider the pattern of configural effects found when the judgment is based either on small differences in spatial frequency or in contrast (Fig. 5, columns 2 and 4). For these judgment types, configural effects were not found when cues to discrimination were presented in different frequency and orientation bands (row 1), nor when presented in different frequency bands at the same orientation (row 4). However, they were found when the two cues to discrimination are presented in the same frequency band at different orientations (rows 2 and 3). These results indicated the existence of a type of summing circuit that is broadband with respect to orientation, but limited with respect to spatial frequency (doughnuts). However, this pattern was specific to spatial frequency and contrast judgments.

Consider the results for judgments of orientations, shown in rows 1 and 4 of column 6 (excepting for the moment the results shown in rows 2 and 3). No configural effects were found when the two cues appear in different frequency and orientation bands (row 1). However, they did appear when the cues appeared in different frequency bands at the same orientation (row 4). These results suggested the existence of a second summing mechanism with different properties than 'doughnuts'. In this case, the mechanism appeared to sum orientation information over a broad range of spatial frequencies, but only within a limited orientation band (cigar).

The results suggested that both contrast and spatial frequency information is carried by mechanisms that sum information over all orientations but only within a limited frequency band. These results might imply that the spatial grain and contrast mechanisms are one and the same. However, other evidence from cue-summation (Thomas & Olzak, 1990), concurrent-judgment (Olzak & Wickens, 1997), uncertainty (Greenlee & Thomas, 1993), and short-term memory experiments (Greenlee, Magnussen & Thomas, 1991) suggests that contrast information is processed independently from information about spatial frequency. Furthermore, the results of one study (Thomas & Olzak, 1990) suggested that when multiple cues to discrimination are presented simultaneously, spatial frequency, contrast and orientation information are all processed independently on one another. These results lend support for our second interpretation of multiple summing mechanisms specialized to signal different types of information about a stimulus, and further suggest that the doughnut profile may actually describe two different mechanisms.

Our analysis does not account for the configural effect found when orientation judgments are made on the second class of stimuli (similar-frequency plaids). We believe this experiment reveals a process quite different from that we have described to account for the other 11 experiments. First, the interaction pattern found when making orientation judgments on similar frequency plaids (Fig. 5, rows 2 and 3) differed from other data sets showing interactions. Although reliable configural effects were found with all types of judgments, masking was considerably reduced in the case of orientation judgments relative to that shown when making spatial frequency or contrast judgments. Second, the quantitative analysis presented in Section 6 of the current paper suggests that the processes underlying the cross-orientation configural effect differs when the

judgment is orientation rather than spatial frequency or contrast. Third, we have mounting evidence that our interpretation of multiple summing mechanisms, specialized to signal abstract information about the orientation, textural grain, and contrast of patterns is correct. In addition to the supporting evidence cited in the previous paragraph, we have recently reported independent evidence for the existence of both cigars and doughnuts in concurrent-judgment (Olzak & Wickens, 1997) and uncertainty (Thomas & Olzak, 1996) experiments. The results of these independent tests support our interpretation of multiple third-stage summing mechanisms, particularly in the case of cigars. We have also made more precise measurements detailing the orientation and spatial frequency profiles of the mechanisms underlying complex spatial frequency, contrast, and orientation judgments (Olzak & Thomas, 1995; Olzak & Thomas, 1996). These profiles agree closely with those we have inferred from our earlier data and in fact serve as the basis for our schematic profiles in Fig. 6.

To the extent that our interpretation of multiple higher-level summing circuits is correct, the small amount of masking and the configural effects found when making orientation judgments on patterns containing similar-frequency, orthogonally oriented components (Fig. 5, rows 2 and 3, columns 5 and 6) still requires an explanation. One possibility is that the output of different 'cigars', which are ideal to signal information about the orientation of object borders, is fed to yet a higher-level mechanism specialized to signal information about the orientation of two-dimensional patterns. The analysis in Section 6 provides some support for this possibility.

6. A test of the model

In previous sections, we argued that the existence and nature of configuration effects provide evidence against direct access models of discrimination and for the existence of summing circuits interposed between primary layer representations and decision processes. In this section we present additional evidence for the existence of summing circuits. The evidence is based on quantitative comparisons of performance in masking conditions with performance in cue summation conditions.

Our model predicts that, when measured by values of d', performance in the same-sign cue summation condition will be the linear sum of performances in the relevant masking conditions, and that performance in the opposite-sign cue summation conditions will be the difference between the performances in the masking conditions. This prediction is most easily seen when q' in Eq. (3) = 1.0. Surprisingly, for the discrimination tasks examined in this paper, the model makes essentially the same prediction (i.e. same within error of measurement) when q' does not equal 1.0. The proof of this fact is given in Appendix A. Thus, the model makes a strong prediction, independent of the value of q', that performance in the same-sign cue-summation conditions will be the sum of performances in the relevant masking conditions, and performance in the opposite sign conditions will be the difference of performances. This section presents the results of quantitative tests that compare predictions of our model to predictions derived from direct-access models. Direct-access models predict that performance in the cue-summation conditions will no better than the root-mean-square of performances in the masking conditions.

We made these tests by calculating expected performance in the two-cue conditions from performance in the two masked conditions. The calculations for each experiment were made on individual data gathered each day in masking and cue-summation conditions. Daily d' values in the two masked conditions were summed to generate expected d' values in the same-sign cue-summation conditions. The absolute value of the difference in performance in masking conditions generated expected d' values in the opposite-sign conditions. Experiments were analyzed separately. For a given experiment, predicted versus obtained values were compared in an analysis of variance procedure, with data from each of the four or five observers participating in the experiment acting as an independent factor. Within-group (within-observer) variability was calculated from the five to eight replications contributed by each observer over the course of the experiment. The data used in our analysis were reported in Thomas & Olzak (1990) and Olzak and Thomas (1991, 1992). We tested data from all experiments that revealed configural effects. These were spatial frequency, contrast, and orientation judgments on similar-frequency plaids (both near 3 or 15 cpd), and orientation judgments on near-vertical gratings of 3 + 15 cpd. The results are shown in Fig. 7.

Fig. 7 plots predicted d' values in the cue summation conditions against observed data for orientation judgments made on vertical components of 3 and 15 cpd (top row), and contrast and spatial frequency judgments made on similar-frequency plaids (second and third rows). Each point represents results from one observer on one day. Different symbols indicate different observers. The diagonal line indicates no difference between predicted and observed values. Points above the line indicate that observed performance was better than predicted; points below the line indicate that observed performance was worse than predicted.

The left-hand column of Fig. 7 illustrates results for same-sign predictions for the three experiments. The results were remarkably uniform across experiments and across observers. Not one of the fourteen data sets tested individually could reject the model prediction,

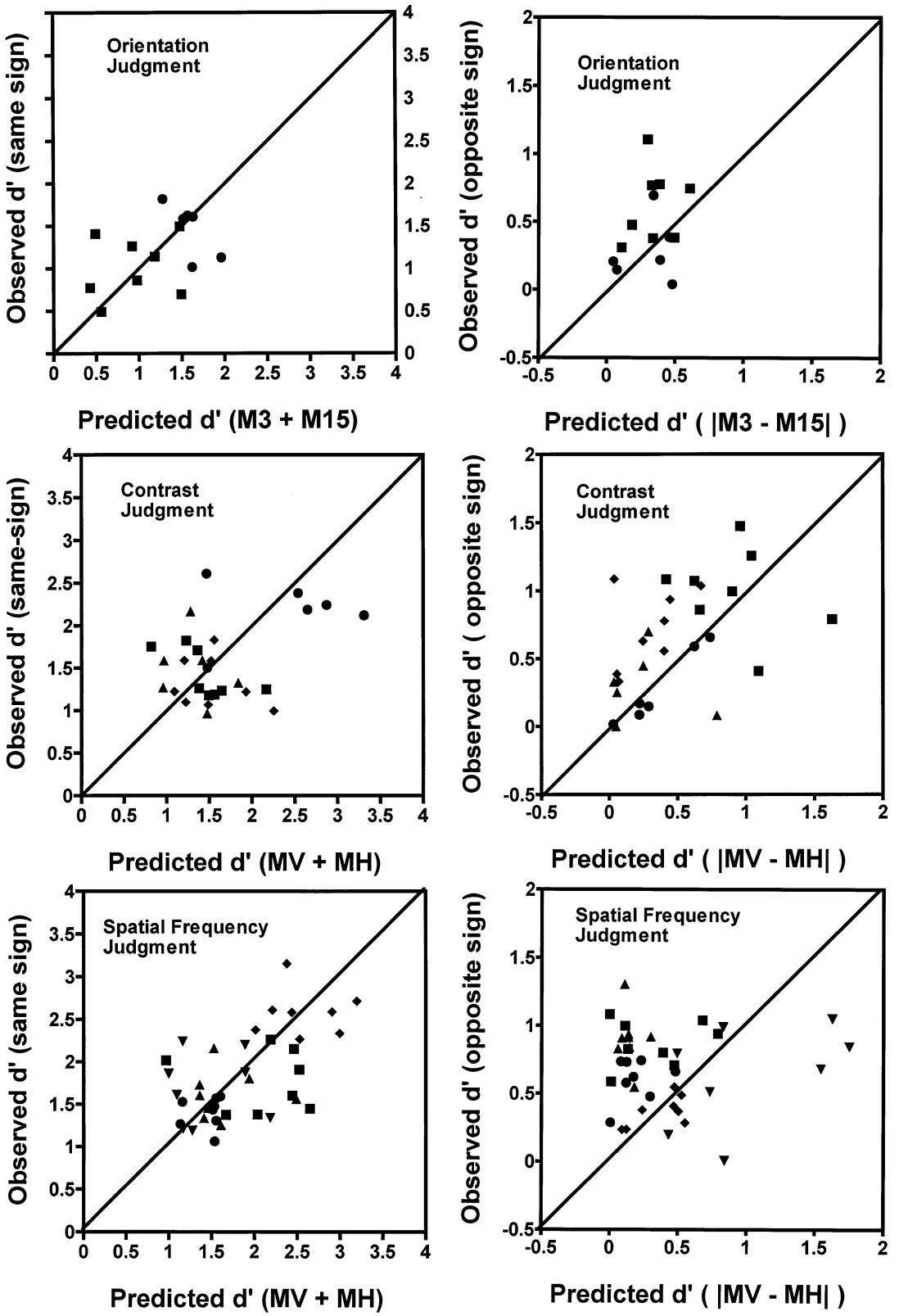

Fig. 7. Predicted vs observed performance in cue-summation analysis for orientation judgments made on two vertical components that differed widely in spatial frequency (top row) and for contrast and spatial frequency judgments made on plaids of similar frequency (middle and bottom rows). The solid diagonal line shows the locus of no difference between predicted and preserved performance. Different symbols represent different observers. In the middle and bottom rows, observers represented by diamonds and triangles were tested with horizontal and vertical components of 15 cpd; remaining symbols represent data collected with 3 cpd components.

nor could group data pooled over observers reject the prediction in any given experiment (all $p > 0.50$).

The data were found to be powerful enough to reject several alternative models. For comparison purposes and to test the power of our data to reject the null hypothesis, we tested an alternative model in which the output of responses to each component are treated as separate and independent sources of information by decision processes. In this information integration model, the relationship between performance in masking conditions (d'_{M1} and d'_{M2})and in the same-sign (d'_{SS}) conditions is given by:

$$d'_{SS} = [d'^{x}_{M1} + d'^{x}_{M2}]^{1/x}$$

The exponent of summation x in the information integration model reflects how information is combined at the decision stage. An exponent of two indicates Euclidean summation and is consistent with ideal-observer integration. Exponents of three to five are generally taken to reflect probability summation. Exponents higher than five represent the use of a single information source. Our data rejected, for every observer individually and as a group in a given experiment, summing exponents of two and higher (all $p < 0.01$). These results lend strong support to the model prediction that performance in the same-sign cue-summation condition will be twice that found in masking condition, and lend strong support for the model as a whole.

The right-hand column of Fig. 7 illustrates results of the analysis for the same experiments and observers in the opposite-sign conditions, which yielded much lower discriminability performance than same-sign conditions (note axes are on different scales in left- and right-hand panels). Here, results are somewhat less uniform across experiments and observers than in the same-sign conditions, although there is some support for model predictions. Neither individual nor group data in two experiments (orientation judgments made on near-vertical 3 and 15 cpd components, top row; and contrast judgments made on similar-frequency plaids, middle row), could reject model predictions (all $p > 0.10$). Group data for these judgments were able to reject predictions of the alternative model. However, we note that performance was quite variable across observers, and in some cases, individual data could not reject the alternative model. The group data of all judgments in the opposite-sign conditions did reject a summing exponent of two (all $p < 0.01$). The data for some individual observers, however, did not have the power to reject summing exponents of two, although higher exponents were always rejected in individual data sets. These results provide some support for the model, although it is not as strong as found in the same-sign cue-summation conditions.

On the other hand, group data obtained when spatial frequency judgments were made on similar frequency plaids in the opposite-sign condition (Fig. 7, bottom right) were found to be inconsistent with model predictions (all $p < 0.001$). Model predictions were also rejected in individual data in four out of five observers. Performance was always better than predicted, and no alternative model was found that fit group data (all $p < 0.01$). Individual data sets were idosyncratically compatible with different models tested, and no meaningful conclusions could be drawn. These results suggest that for spatial frequency judgments on similar frequency plaids, observers are not using only the transformed output of third-stage summing mechanisms when components are combined in opposition, but may have access to other information.

Several interpretations of this last set of findings are possible. One alternative, consistent with the findings of Thomas and Olzak (1997) with respect to orientation tuning of the gain control pool, is that 'doughnuts' do not sum all orientations equally. Instead they may have a weighting function that declines with increasing differences in orientation. The higher-than-predicted performance in the opposite-sign conditions might then result from probability summation over 'doughnuts' that prefer different orientations.

A second possibility is that observers may have some access back down to the primary layer, and use this information (possibly blended with responses from the output of summing circuits) to make discrimination decisions in the absence of a useful summing circuit. Responses from primary layer mechanisms presumably would be diminished by the operation of nonlinear gain control processes (as in masking) and would not be expected to lead to high levels of performance, but would improve performance relative to that expected if observers only use the output of summing circuits upon which to base their decision. A third possibility is that observers are able to use shape or other local cues in the stimulus to perform the discrimination, albeit not efficiently. Each of these hypotheses is consistent with additional data gathered in independent concurrent-response, uncertainty, and discrimination studies in the case of frequency judgments on plaids (Thomas, Olzak & Shimozaki, 1992; Olzak, Wickens & Thomas, 1993a,b; Thomas & Olzak, 1996; Olzak & Wickens, 1997). The results of the additional studies agree that frequency judgments made on plaids are only partially mediated by a summing circuit, but do not distinguish among the alternative interpretations. Similar tests made with $3 + 15$ vertical components, however, strongly support the notion that orientation discriminations made on these components are mediated solely by a summing circuit such as the 'cigar' we have described.

Fig. 8 shows results of the analysis for orientation judgments on similar-frequency plaids. This is the case we identified as showing a pattern of results inconsistent with the other results, notably showing only minimal masking.

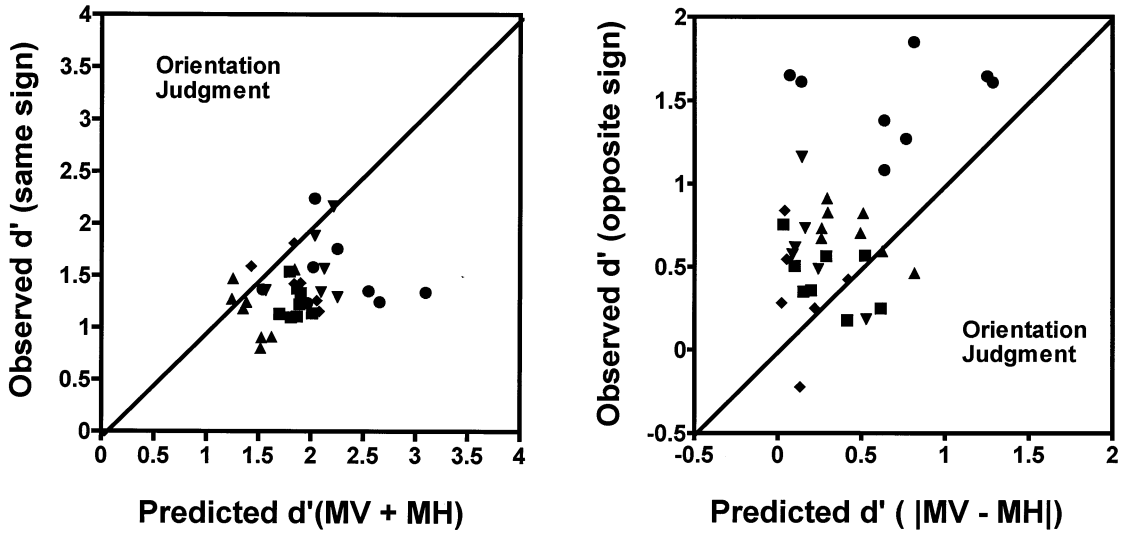

Fig. 8. Predicted vs observed performance in cue-summation analysis for orientation judgments made on similar-frequency plaids. Data in each panel is plotted as in Fig. 7.

Panel A shows results of the analysis for results obtained with vertical and horizontal components of similar frequency in the same-sign cue-summation condition (data from both the $3+3$ cpd and the $15+15$ cpd experiments are intermingled). Panel B shows performance in the opposite-sign condition. Here, we demonstrate that unlike the data obtained in the experiments shown in Fig. 7, these data reject the notion that a summing circuit combines responses prior to the decision stage in both the same- and opposite-sign cases (both $P < 0.001$). The results are more consistent with the alternative model that combination occurs at the decision stage of processing ($x = 2$ ($P > 0.500$) or $x = 3$ ($P > 0.137$)), although no single alternative model was able to fit all of the data of individuals.

This pattern of results differs considerably from the other experiments showing configural effects and suggest that these judgments are mediated by mechanisms with different characteristics than those mediating the other judgments. One possibility is that these judgments are mediated by yet a higher-level mechanism, which takes input from the orientation-signalling mechanism. This hypothesis is currently under test.

7. Discussion

In this paper, we have presented a general framework within which we can quantitatively test hypotheses about how the output from linear filters at the initial layer of cortical processes are transformed at higher levels of processing when performing fine spatial discriminations. The general multistage form of the model was derived in part from other hierarchical models in the literature (see Section 1), but detailed characterizations at each processing stage reflect empirical findings from our cross-orientation and cross-frequency masking studies and cue-summation studies. In brief, we have posited that responses of early cortical filters are nonlinearly transformed both by independent within-pathway processes and by a divisive gain control process that includes filters with a broad range of orientation tuning characteristics, but limited with respect to spatial frequency (Eq. (2)). We have further proposed that specific subsets of these transformed responses are subsequently combined by multiple higher-level summing circuits that are specialized to provide information about particular aspects of stimuli that are complex in the Fourier domain (Eq. (3)).

Our findings thus far have led us to propose two families of higher-level mechanisms. One family, colloquially termed 'doughnuts' to describe their profiles in the Fourier plane, sums over a broad range of orientations, but only within limited frequency ranges. The existence of this type of mechanism is revealed when discriminations are based on spatial frequency or contrast information, and appears specialized to signal spatial grain and/or information about the contrast of textured surfaces. The second family, colloquially termed 'cigars', sums over a broad range of spatial frequencies, but only within limited orientation ranges. This type of mechanism is revealed when discrimination judgments are based on orientation, and appears to be specialized to signal information about the orientation of edges, object borders, and textures.

The normalization portion of the model as it stands accounts for much, but not all of the data from which it was derived and against which it has been tested. An important question that remains open is whether the single gain-control pool with weighted inputs from

primary-level units described in Eq. (2) is adequate as it stands, or whether some gain control processes are associated with higher-level summing circuits, as our masking data from the experiment with $3+15$ cpd components and orientation judgments might suggest. We note that a single-pool model is more parsimonious and should not be rejected without good cause, and therefore reserve judgment on this issue until further tests can be made.

One limitation of the gain control processes as we have described them is that they cannot fully account for contrast discrimination data (Legge & Foley, 1980; Smith & Thomas, 1989; Bowne, 1990). As described, the gain control processes cause the responses of individual pathways and summing circuits to saturate and become independent of contrast at medium and high contrasts. This putative saturation is consistent with the observed fact that orientation and spatial frequency discrimination performance also becomes independent of contrast at these levels, but falsely predicts that contrast discrimination will disappear. Legge and Foley (1980) and Foley (1994) address this problem by making their analogue of the exponent p in Eq. (2) take a slightly larger value in the numerator(s) than in the denominator(s). With this modification, the responses of pathways and summing circuits do not fully saturate at any contrast, thus preserving the possibility of contrast discrimination. However, this modification alone is not wholly satisfactory because it sacrifices the ability to account for the saturation of orientation and spatial frequency discrimination at medium and high contrasts. A possible solution to the dilemma is if the noise which perturbs the recoding and decision processes increases with stimulus contrast, then orientation and spatial frequency discrimination will saturate at medium and high contrasts even with the Legge and Foley modification (Thomas, 1983). Geisler and Albrecht (1997) found physiological evidence of such signal dependent noise, but Smith and Thomas (1989) were unable to find psychophysical evidence for it in suprathreshold contrast discrimination tasks. Teo and Heeger (1994) proposed an entirely different answer to the dilemma, suggesting that pathways are tuned to different contrast ranges, such that when one pathway saturates and can no longer mediate discrimination, another pathway, tuned to a higher range, takes over.

A question that has been raised about hierarchical models in general, such as those discussed in Section 1, is whether the mechanisms isolated are peculiar to the stimuli and tasks of particular experiments rather than representing general early recoding mechanisms of spatial vision. Our work includes the development of a battery of discrimination tasks (masking, cue-summation, uncertainty experiments and concurrent-response studies) that isolate different stages of processing and mechanisms within the context of the model. These permit us to test our general framework in different experimental contexts and to modify the model as required by consistent discrepancies between model predictions and the data. As we have noted in earlier sections of this paper, work described elsewhere (Thomas & Olzak, 1996; Olzak & Wickens, 1997), has demonstrated results consistent with our model in two additional paradigms: an uncertainty paradigm in which observers must monitor two potential cue sources, only one of which is informative on any given trial, and a concurrent-response paradigm in which simultaneous judgments are made about the identity of two components. These data converge with the single-response discrimination results analyzed here to strongly suggest that orientation judgments on stimuli that are complex in the Fourier domain are mediated solely by 'cigars'. Similarly, the results agree that spatial frequency judgments made on texturelike patterns (similar frequency, multiple orientations) are mediated by 'doughnuts', although in this case other information may also be used. In both instances, results are consistent with the existence of the third-stage summing circuits we have described and are entirely unexpected from known properties of first-layer mechanisms.

One consequence of the recoding scheme we have proposed is that it provides an explanation of how different types of sensory information become independently represented at higher levels of processing. Along with other types of information, the stimulus dimensions of spatial frequency, orientation, and contrast information are jointly coded at the primary cortical layer of processing and therefore entirely confounded. For example, at the primary level of representation, the differential response to two slightly different spatial frequencies or orientations can be identical to changing the contrast of one of the frequencies. Representing different types of information via separate higher-level mechanisms that the observer can choose from to use as the basis for a discrimination decision solves this confounding problem. A wealth of evidence suggests that decisions about different stimulus dimensions are made independently (Graham, 1989; Ashby, 1992).

A question that remains unanswered by the data analyzed here is whether a single family of 'doughnut' mechanisms signal both spatial frequency and contrast information, or whether the information about the two dimensions is transmitted via separate families with similar profiles in the Fourier domain. Several studies have addressed the question of how contrast and spatial frequency information are represented when making discrimination judgments, and the evidence is consistent with the notion of two separate and independent representations for the two types of information. First, discrimination thresholds increase when an observer is uncertain as to whether the cue to discrimination is a contrast or a spatial frequency cue. This uncertainty

effect only occurs when processing of two cues is performed by two separate mechanisms. The magnitude of the increase in these experiments quantitatively agrees with that predicted by a model that assumes independent processing (Greenlee & Thomas, 1993). Second, contrast discrimination performance declines when the interval between two successively-presented stimuli increases. No such decline is found when the discrimination is based on spatial frequency (Greenlee et al., 1991). Finally, discrimination performance increases over single-cues conditions when both contrast and spatial frequency cues to discrimination are available, by an amount indicating Euclidean summation of information (Thomas & Olzak, 1990). Although the latter result does not necessarily imply separate and independent pathways, it is consistent with that interpretation and with conclusions of the other studies.

The neural substrates of the summing circuits we describe are entirely unknown, although it is tempting to draw parallels between the psychophysics and physiological findings. In both monkey and cat, striate cells tuned to many different spatial frequencies are anatomically organized in highly specific orientation columns through all six cortical layers (Hubel, Wiesel & Stryker, 1978; DeValois & DeValois, 1988). In 2-deoxyglucose experiments, long cortical slabs appear in response to a stimulus containing many frequencies at a single orientation. Summation within these columns or slabs might well provide the basis for the summing 'cigars' we describe. The summed information may be represented in V1 itself, or may be represented in V2 or beyond. Peterhans and von der Heydt (1993) have described single cells in V2 with response properties not unlike our 'cigars'. Like the units we describe, these cells appear specialized to signal information about stimuli that contain a broad range of frequencies within a narrow orientation.

The units described by Peterhans and von der Heydt may be implicated in other tasks that may require fine orientation contour integration over gaps in space (Field et al., 1993), the perception of illusory contours (Dresp & Bonnet, 1995), long-range facilitation and suppression of contrast (Polat & Sagi, 1993, 1994a,b; Polat & Norcia, 1996; Zenger & Sagi, 1996), and general object mechanisms that signal object borders (Grossberg & Mingolla, 1985a,b; Shapley & Gordon, 1985). Similar units have also been proposed to mediate alignment tasks, bisection, and long-range separation discrimination (Hess & Badcock, 1995; Waugh & Levi, 1995; Levi & Waugh, 1996; Mussap & Levi, 1997). Although we have not yet investigated the spatial properties of our orientation-signaling mechanisms nor whether they integrate information over gaps in space, the possibility remains that mechanisms proposed to mediate these tasks and the 'cigars' we have described are in actuality the same mechanisms. The possibility that the third-stage units we describe play a role in some of these tasks is strengthened by preliminary work we have reported suggesting that the units are very narrowly tuned with respect to orientation (Olzak & Thomas, 1995, 1996) and are not sensitive to phase (Stankiewicz et al., 1995), characteristics consistent with those reported for units mediating the tasks described above. An obvious next step is to extend our investigations into the spatial domain to determine how the mechanisms we have isolated normalize and integrate information over space.

The physiological substrate underlying summation of frequency and contrast information over orientations is also unknown. However, long cortical slabs also appear in 2-deoxyglucose experiments when the stimulus is a single spatial frequency at all orientations (Hubel et al., 1978; Tootell, Silverman & DeValois, 1981; Silverman, 1984). It is possible that this organization provides the basis for summation within 'doughnuts' when making spatial frequency (or contrast) judgments. The mechanism mediating orientation judgments on plaids is the least well-characterized in our data; thus, it is difficult to know what neural properties might be important. However, we note that neurons in V4 of primate visual cortex which respond preferentially to cross-like stimuli have been described by Gallant et al. (1993). Although these could not serve as general summing circuits over many orientations, they could potentially signal 2-dimensional orientation information about plaid or crosslike stimuli.

In closing, we note that the third-stage mechanisms we have isolated in the context of the model are elegantly tailored to perform real-world tasks requiring fine spatial discriminations. Real world features are complex in the Fourier domain, stimulating many different sets of V1 neurons. The two complementary mechanism families we describe maximize the ability to discriminate differences in the orientation of real edges or object borders from any viewing distance, or the spatial grain (or contrast) of a surface seen at any orientation; the former is size-invariant; the latter orientation-invariant. Furthermore, the sign-preserving summation we observe in both second order mechanisms favors the ability to distinguish among complex patterns and stimulus features whose Fourier components have all undergone the same rotational or in-depth translation over those whose components rotate or translate in different directions. That is, they respond best to the way real-world stimulus features translate or rotate as rigid bodies. These properties make our mechanisms likely candidates to play significant roles in everyday visual tasks requiring fine spatial discriminations.

Acknowledgements

The authors wish to thank Thomas D. Wickens for numerous theoretical and mathematical discussions while developing the model, for developing Appendix A as an analytic demonstration of model predictions, for significant contributions in developing the mathematical notation we use throughout the paper, and for help in the preparation of several figures. We are also very grateful to Norma Graham for ongoing discussions and for an extraordinary review of an earlier version of this manuscript. Finally, we thank Charles Chubb for a thoughtful review and valuable discussions. This research was supported by NIH grant EY360 to JPT and LAO. An initial version of the model was presented at ARVO in 1992.

Appendix A. The relationship between d' for masked and configural conditions

Let A and B be discriminative gratings and M the intermediate mask grating. In the masked task, the observer discriminates between the stimuli (A + M) and (B + M) (the first component is vertical, the second horizontal); in the same-sign configural task, the stimuli are (A + A) and (B + B); in the opposite-sign configural task, they are (A + B) and (B + A). Assume the model presented in Eqs. (1–5), and make the following simplifying assumptions:

1. Separate pathways. The horizontal and vertical components activate separate pathways at the primary layer that do not interact.
2. Component equivalence. The summing-circuit weights $v_\alpha(i)$ in Equation (3) are the same for the two components stimulus (horizontal and vertical or 3 cpd and 15 cpd, depending on condition).
3. Weight stability. The summing-circuit weights $v_\alpha(i)$ do not vary across experimental conditions.
4. Small differences. The components A, B, and M are closely spaced. Consequently, (a) equal gain control: the three components activate the same gain-control pool; (b) local linearity: the response function can be treated as linear for these components.

Denote the response of unit $R_i^{(j)}$ to stimulus X by $R_i^{(j)}(\mathrm{X})$, and, for notational simplicity, use q for the third-stage exponent in Eq. (3) instead of q'. Because gain control is the same for all stimuli (assumption 4 a), the second stage response $R_i^{(2)}(\mathrm{X})$ depends only on $R_i^{(1)}(\mathrm{X})$, not on the other components of the stimulus. By the separate-pathways assumption (assumption 1), the third-stage response to a two-component stimulus separates into that due to each of its components. In a third-stage unit tuned to A, the response to the masked stimulus (A + M) is:

$$R_{\mathrm{A}}^{(3)}(\mathrm{A+M}) = \left\{\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{A})]^q + \sum v_{\mathrm{A}}(i)[R_i^{(2)}(M)]^q\right\}^{1/q}$$
$$= K'\left\{1 + \frac{1}{K}\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{A})]^q\right\}^{1/q},$$

where $K = \Sigma\ v_{\mathrm{A}}(i)[R_i^{(2)}(M)]^q$ is the response to the mask and $K' = K^{1/q}$. Denote this response by $K'(1+T)^{1/q}$, where T is the ratio of the response of a third-stage unit to its appropriate stimulus component to its response to the mask. Similarly, the response of this unit to the other stimulus is:

$$R_{\mathrm{A}}^{(3)}(\mathrm{B+M}) = \left\{\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{B})]^q + \sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{M})]^q\right\}^{1/q}$$
$$= K'\left\{1 + \frac{1}{K}\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{B})]^q\right\}^{1/q}$$
$$= K'(1+t)^{1/q},$$

where t is the ratio of the response to the other component to the response to the mask. Similar derivations, using the assumption of component equivalence (assumption 2), give the responses to the configural stimuli. For the preferred stimulus in the same-sign task,

$$R_{\mathrm{A}}^{(3)}(\mathrm{A+A}) = \left\{\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{A})]^q + \sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{A})]^q\right\}^{1/q}$$
$$= K'\left\{1 + \frac{1}{K}\sum v_{\mathrm{A}}(i)[R_i^{(2)}(\mathrm{A})]^q\right\}^{1/q}$$
$$= K'(2T)^{1/q}.$$

For the non-preferred stimulus, the analogous calculation gives $R_{\mathrm{A}}^{(3)}(\mathrm{B+B}) = K'(2t)^{1/q}$. By the weight-stability assumption and local linearity (assumptions 3 and 4b), a similar analysis of the unit tuned to component B gives results with the same values of T and t:

$$R_{\mathrm{B}}^{(3)}(\mathrm{B+M}) = K'(1+T)^{1/q},$$
$$R_{\mathrm{B}}^{(3)}(\mathrm{A+M}) = K'(1+t)^{1/q},$$
$$R_{\mathrm{B}}^{(3)}(\mathrm{B+B}) = K'(2T)^{1/q},$$
$$R_{\mathrm{B}}^{(3)}(\mathrm{A+A}) = K'(2t)^{1/q},$$

Performance measures for the two conditions are found using Equations (4) and (5):

$$d'_{\mathrm{mask}} = \frac{[R_{\mathrm{A}}^{(3)}(\mathrm{A+M}) - R_{\mathrm{B}}^{(3)}(\mathrm{A+M})]}{-[R_{\mathrm{A}}^{(3)}(\mathrm{B+M}) - R_{\mathrm{B}}^{(3)}(\mathrm{B+M})]}{\sigma}$$
$$= \frac{2K'}{\sigma}\,[1+T]^{1/q} - (1+t)^{1/q}$$

$$d'_{\mathrm{SS\ config}} = \frac{[R_{\mathrm{A}}^{(3)}(\mathrm{A+A}) - R_{\mathrm{B}}^{(3)}(\mathrm{A+A})]}{-[R_{\mathrm{A}}^{(3)}(\mathrm{B+B}) - R_{\mathrm{B}}^{(3)}(\mathrm{B+B})]}{\sigma}$$
$$= \frac{2K'}{\sigma}\,\{[2T]^{1/q} - (2t)^{1/q}\}$$

Using local linearity (assumption 4b), the response to the mask is intermediate between the response to A and B, so that T and t are symmetric about one: for some small positive ε, write $T = 1 + \varepsilon$ and $t = 1 - \varepsilon$. With this substitution, the ratio of the d' statistics for the two conditions is:

$$r(q,\varepsilon) = \frac{d'_{\text{SS config}}}{d'_{\text{mask}}} = \frac{(2+2\varepsilon)^{1/q} - (2-2\varepsilon)^{1/q}}{(2+2\varepsilon)^{1/q} - (2-2\varepsilon)^{1/q}}.$$

Although this ratio depends on q and ε, it is very close to two for reasonable values of these parameters. Specifically, $r(q, \varepsilon)$ exactly equals two for any ε when $q = 1$ or in the limit as $\varepsilon \to 0$ for any q. The partial derivative of $r(q, \varepsilon)$ with respect to either argument is positive for positive arguments; thus for any $0 < q < q_{\max}$ and $0 < \varepsilon < \varepsilon_{\max}$, $r(q, \varepsilon)$ lies between two and $r(q_{\max}, \varepsilon_{\max})$. When $q = 50$ and $\varepsilon = 0.2$ (i.e. the response to the unit's preferred component is 150% of that to the non-preferred component), $r(q, \varepsilon)$ differs from two by less than 1%. For smaller (and more plausible) arguments, the ratio is still closer to two, well less than can be discriminated by these experiments. We conclude that a 2:1 ratio cannot be taken as evidence for linear summation.

A similar analysis applied to the opposite-sign configural condition shows that when the horizontal and vertical components are equally discriminable, the responses of symmetrically-placed units to the stimulus a

$$R_{\text{A}}^{(3)}(\text{A} + \text{B}) = R_{\text{B}}^{(3)}(\text{A} + \text{B})$$

and

$$R_{\text{A}}^{(3)}(\text{B} + \text{A}) = R_{\text{B}}^{(3)}(\text{B} + \text{A}).$$

With these identities, the model predicts that $d'_{\text{OSconfig}} = 0$.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*, 284–299.

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity & the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, *7*, 531–546.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, *48*, 217–237.

Ashby, F. G. (1992). *Multidimensional models of perception and cognition*. New Jersey: Lawrence Erlbaum Associates.

Bergen, J. R., & Adelson, E. H. (1986). Visual texture segmentation based on energy measures. *Journal of the Optical Society of America A*, *3*, 98.

Bergen, J. R., & Landy, M. S. (1991). Computational modeling of visual texture segregation. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing*. Cambridge, MA: MIT Press, 291–301.

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, *2*, 41–55.

Bovik, A. C., Clark, M., & Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Trans Pattern Anal Mach Intell*, *12*, 55–73.

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation, or temporal frequency discrimination. *Vision Research*, *30*, 449–461.

Boynton, R. M., Ikeda, M., & Stiles, W. S. (1964). Interactions among chromatic mechanisms as inferred from positive and negative increment thresholds. *Vision Research*, *4*, 87–177.

Burbeck, C. A., & Yap, Y. L. (1990a). Spatial-filter selection in large-scale spatial-interval discrimination. *Vision Research*, *30*, 263–272.

Burbeck, C. A., & Yap, Y. L. (1990b). Two mechanisms for localization? Evidence for separation-dependent and separation-independent processing of position information. *Vision Research*, *30*, 739–750.

Burbeck, C. A. (1991). Encoding Spatial Relations. In R. J. Watt, *Pattern recognition by man and machine*. Boca Raton: CRC Press.

Burr D.C., & Morrone M.C. (1994) Features in visual images. Higher order processing in the visual system. Ciba Foundation Symposium 184, 129–146. New York: Wiley.

Cannon, M. W., & Fullenkamp, S. C. (1991). A transducer model for contrast perception. *Vision Research*, *31*, 983–988.

Cannon, M. W., & Fullenkamp, S. C. (1996). A model for inhibitory lateral interaction effects in perceived contrast. *Vision Research*, *36*, 1115–1125.

Chubb, C., & Landy, M.S. (1991) Orthogonal distribution analysis: A new approach to the study of texture perception. In M.S. Landy, & J.A. Movshon, *Computational models of visual processing*, pp. 291–301. Cambridge, MA: MIT Press.

Chubb, C., McGowan, J., Sperling, G., & Werkhoven, P. (1994). *Non-Fourier motion analysis. Higher order processing in the visual system*. (Ciba Foundation Symposium 184). New York: Wiley.

Cohn, T. E., & Lasley, D. J. (1976). Binocular vision: Two possible central interactions between signals from the two eyes. *Science*, *192*, 561–562.

Debruyn, E. J., & Bonds, A. B. (1986). Contrast adaptation in cat visual cortex is not mediated by GABA. *Brain Research*, *383*, 339–342.

Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, *25*, 1869–1978.

Derrington, A. M., & Henning, G. B. (1994). Implications of motion detection for early non-linearities. Ciba Foundation Symposium 184. 211–20; *discussion*, *220–6*, 269–271.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, *11*(5), 219–226.

Dresp, B., & Bonnet, C. (1995). Subthreshold summation with illusory contours. *Vision Research*, *35*, 1071–1078.

Enroth-Cugell, C., & Robson, J. G. (1984). Functional characteristics and diversity of cat retinal ganglion cells. *Investigative Ophthalmology and Visual Science*, *25*, 250–267.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research*, *33*, 173–193.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A*, *6*, 1710–1719.

Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science*, *259*, 100–103.

Geisler, W. S., & Albrecht, D. G. (1995). Bayesian analysis of identification performance in monkey visual cortex: nonlinear mechanisms and stimulus certainty. *Vision Research*, *35*, 2723–2730.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monekeys and cats: Detection, discrimination, and identification. *Visual Neuroscience*, *14*, 897–919.

Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London Series B Biological Science*, *249*, 235–245.

Georgeson, M. A. (1994). *From filters to features*: *location*, *orientation*, *contrast*, *and blur*. *Higher order processing in the visual system*. (Ciba Foundation Symposium 184). New York: Wiley, 147–164.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel andmultiple-channels models. *Vision Research*, *11*, 251–259.

Graham, N., & Sutter, A. (1996). Effect of spatial scale and background luminance on the intensive and spatial nonlinearities in texture segregation. *Vision Research*, *36*, 1371–1390.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier and complex (non-Fourier) texture channels. *Vision Research*, *38*, 231–257.

Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research*, *32*, 719–743.

Greenlee, M. W., & Thomas, J. P. (1993). Simultaneous discrimination of the spatial frequency and contrast of periodic stimuli. *Journal of the Optical Society of America A*, *10*, 395–404.

Greenlee, M. W., Magnussen, S., & Thomas, J. P. (1991). Different neural codes for spatial frequency and contrast. In A. Valberg, & B. B. Lee, *From Pigments to Perception*. New York: Plenum Press.

Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, *166*, 1303–1306.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychology Review*, *92*, 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: texture, boundaries, and emergent segmentation. *Perception and Psychophysics*, *38*, 141–171.

Heeger, D. J. (1991). Computational model of cat striate physiology. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing*. Cambridge, MA: MIT Press, 291–301.

Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *Journal of Neurophysiology*, *70*, 1885–1898.

Hess, R. F., & Badcock, D. R. (1995). Metric for separation discrimination by the human visual system. *Journal of the Optical Society of America A*, *12*, 3–16.

Hess, R. F., & Field, D. J. (1995). Contour integration across depth. *Vision Research*, *35*, 1699–1711.

Hess, R. F., & Hayes, A. (1993). Neural recruitment explains "Weber's law" of spatialposition. *Vision Research*, *33*, 1673–1684.

Hess, R. F., & Hayes, A. (1994). The coding of spatial position by the human visual system: effects of scale and retinal eccentricity. *Vision Research*, *34*, 625–643.

Hess, R. F., & Holliday, I. E. (1992). The coding of spatial position by the human visual system: Effects of spatial scale and contrast. *Vision Research*, *32*, 1085–2097.

Hubel, D. H., Wiesel, T. N., & Stryker, M. P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *Journal of Computing and Neurology*, *177*, 361–380.

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1.0 second: Theoretical predictions and empirical validation. *Journal of the Optical Society of America A*, *2*, 1170–1190.

Klein, S. A., Stromeyer, C. F., & Ganz, L. (1974). The simultaneous spatial frequency shift: A dissociation between the detection and perception of gratings. *Vision Research*, *14*, 1421–1432.

Laurinen, P. I., Olzak, L. A., & Peromaa, T. L. (1997). Early cortical influences in object segregation and the perception of surface lightness. *Psychological Science*, *8*, 386–390.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, *70*, 1458–1471.

Lennie, P. (1980). Parallel visual pathways: A review. *Vision Research*, *20*, 561–594.

Levi, D. M., & Waugh, S. J. (1996). Position acuity with opposite-contrast polarity features: Evidence for a nonlinear collector mechanism for position acuity. *Vision Research*, *36*, 573–588.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Academy of London B*, *200*, 269–294.

Meese, T. S. W., & Georgeson, M. A. (1996a). The tilt aftereffect in plaids and gratings: channel codes, local signs, and "patchwise transforms". *Vision Research*, *36*, 1421–1437.

Meese, T. S. W., & Georgeson, M. A. (1996b). Spatial filter combination in human pattern vision: channel interactions revealed by adaptation. *Perception*, *25*, 255–277.

Meese, T. S. (1995). Phase-reversal discrimination in one and two dimensions: performance is limited by spatial repetition, not by spatial frequency content. *Vision Research*, *35*, 2157–2167.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision andspatial vision: Two cortical pathways. *Trends in Neuroscience*, *6*, 414–417.

Morgan, M. J., & Hotopf, W. H. N. (1989). Perceived diagonals in grids andlattices. *Vision Research*, *29*, 1005–1015.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase-dependent energy model. *Proceedings of the Royal Academy of London B Biological Science*, *235*, 221–245.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. *Proceedings of the Royal Society of London Biology*, *216*, 335–354.

Moulden, B. (1994). *Collator units*: *second-stage orientational filters*. *Higher Order Processing in the Visual System*. (Ciba Foundation Symposium 184). New York: Wiley, 170–184.

Mussap, A. J., & Levi, D. M. (1997). Vernier acuity with plaid masks: The role of oriented filters in vernier acuity. *Vision Research*, *37*, 1325–1340.

Nielsen, K. R. K., Watson, A. B., & Ahumada, A. J. (1985). Application of a computable model of human spatial vision to phase discrimination. *Journal of the Optical Society of America A*, *2*, 1600–1606.

Olzak, L. A., & Thomas, J. P. (1981). Gratings: why frequency discrimination is sometimes better than detection. *Journal of the Optical Society of America*, *71*, 64–70.

Olzak, L. A., & Thomas, J. P. (1986). Seeing spatial patterns. In K. R. Boff, L. Kaufman, & J. P. Thomas, *Handbook of perception and human performance*. New York: Wiley, 7.1–7.56.

Olzak, L. A., & Thomas, J. P. (1991). When orthogonal orientations are not processed independently. *Vision Research*, *31*, 51–57.

Olzak, L. A., & Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision Research*, *32*, 1885–1898.

Olzak, L. A., & Thomas, J. P. (1995). Normalizing and summing pools in complex orientation discriminations. *Supplementary Optics and Photonics News*, *6*(7), 11–12.

Olzak, L. A., & Thomas, J. P. (1996). Characteristics of third-stage mechanisms in pattern vision. *Supplement Investigative Ophthalmology and Visual Science*, *37*, 1148.

Olzak, L. A., & Wickens, T. D. (1997). Complex pattern discrimination: Orientation information is integrated across spatial scale; spatial frequency and contrast information is not. *Perception*, *26*, 1101–1120.

Olzak, L. A., Wickens, T. D., & Thomas, J. P. (1993a). Why we can't see the forest for the trees: Serial processing of disparate spatial frequency bands. *Supplement Perception*, *22*, 5.

Olzak, L. A., Wickens, T. D., & Thomas, J. P. (1993b). Simultaneously discriminating different aspects of complex patterns: When you can, when you can't and why. OSA Annual Meeting Technical Digest. *Opt Soc Am*, *16*, 267–268.

Peterhans, E., & von der Heydt, R. (1991). Bridging the gap between psychophysics and physiology. *Trends in Neuroscience*, *14*(3), 112–119.

Peterhans, E., & von der Heydt, R. (1993). Functional organization of area V2 in the alert macaque. *European Journal of Neuroscience*, *5*, 509–524.

Polat, U., & Norcia, A. M. (1996). Neurophysiological evidence for contrast dependent long–range facilitation and suppression in the human visual cortex. *Vision Research*, *36*, 2099–2109.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, *33*, 993–999.

Polat, U., & Sagi, D. (1994a). Spatial interactions in human vision: from near to far via experience-dependent cascades of connections. *Proceedings of the National Academy of Science USA*, *91*(4), 1206–1209.

Polat, U., & Sagi, D. (1994b). The architecture of perceptual spatial interactions. *Vision Research*, *34*, 73–78.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris, *Visual coding and adaptability*. Hillsdale, NJ: Erlbaum, 177–214.

Sagi D. (1991). Spatial filters in texture segmentation tasks. In D. Blum, *Channels in the visual nervous system: Neurophysiology, Psychophysics, and Models*, (pp. 397–424), London: Freund Publishers.

Shapley, R., & Gordon, J. (1985). Nonlinearity in the perception of form. *Perception and Psychophysics*, *37*, 84–88.

Shapley, R. M., & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual Review of Neuroscience*, *8*, 547–583.

Silverman, M.S. (1984). *Deoxyglucose and electrophysiological evidence for spatial frequency columns in cat striate cortex*. Doctoral Dissertation, University of California, San Francisco.

Smith, B. G., & Thomas, J. P. (1989). Why are some spatial discriminations independent of contrast? *Journal of the Optical Society of America A*, *6*, 713–724.

Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, *34*, 2239–2257.

Solomon, J. A., & Sperling, G. (1995). 1st- and 2nd-order motion andtexture resolution in central and peripheral vision. *Vision Research*, *35*, 59–64.

Sperling, G., Chubb, C., Solomon, J. A., & Lu, Z. L. (1994). *Full-wave and half-wave rectification in motion and texture. Higher order processing in the visual system*. (Ciba Foundation Symposium 184). New York: Wiley, 193–205.

Stankiewicz, B. J., Thomas, J. P., & Olzak, L. A. (1995). Multiple component orientation discrimination invariant with respect to phase. Paper presented at the Optical Society of America Meeting, Portland, OR. *Supplementary Optical and Photonics News*, *6*(7), 139.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science*, *262*, 685–688.

Teo, P., & Heeger, D. (1994). *Perceptual Image Distortion First IEEE Int Confer Image Process*, *2*, 982–986.

Thomas, J. P., & Gille, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America*, *69*, 652–660.

Thomas, J. P., & Olzak, L. A. (1990). Cue summation in spatial discriminations. *Vision Research*, *30*, 865–875.

Thomas, J. P., & Olzak, L. A. (1996). Uncertainty experiments support the roles of second-order mechanisms in spatial frequency and orientation discriminations. *Journal of the Optical Society of America A*, *13*, 689–696.

Thomas, J. P., & Olzak, L. A. (1997). Contrast gain control and fine spatial discriminations. *Journal of the Optical Society of America A*, *14*, 2392–2405.

Thomas, J. P., & Shimamura, K. K. (1975). Inhibitory interaction between visual pathways tuned to different orientations. *Vision Research*, *15*, 1373–1380.

Thomas, J. P. (1970). Model of the function of receptive fields in human vision. *Psychology Review*, *77*, 121–134.

Thomas, J. P. (1983). Underlying psychometric function for detecting gratings and identifying spatial frequency. *Journal of the Opt Society of America*, *73*, 751–758.

Thomas, J. P. (1985). Detection and identification: How are they related? *Journal of the Optical Society of America A*, *2*, 1457–1467.

Thomas, J. P. (1989). Independent processing of suprathreshold spatial gratings as a function of their separation in spatial frequency. *Journal of the Optical Society of America A*, *2*, 1457–1467.

Thomas, J. P., Olzak, L. A., & Shimozaki, S. S. (1992). Using distinctive Fourier components to discriminate between complex patterns. *Ophthalmology Physiology Optics*, *12*, 189–192.

Tootell, R. B. H., Silverman, M. S., & DeValois, R. L. (1981). Spatial frequency columns in primary visual cortex. *Science*, *214*, 813–815.

Van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neuroscience*, *6*, 370–375.

Van Essen, D. C. (1985). Functional organization of primate visual cortex. In E. G. Jones, & A. Peters, *Cerebral cortex*. New York: Plenum, 3259–3329.

Wang, H., & Levi, D. M. (1994). Spatial integration in position acuity. *Vision Research*, *34*, 2859–2877.

Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research*, *21*, 1115–1122.

Watson, A. B. (1983). Detection and recognition of simple spatial forms. In O. J. Braddick, & A. C. Sleigh, *Physiological and biological preprocessing of images*. New York: Springer, 110–114.

Watt, R. J., & Morgan, M. J. (1984). Spatial filters and the localization of luminance changes in human vision. *Vision Research*, *24*, 1387–1397.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, *25*, 1661–1674.

Waugh, S. J., & Levi, D. M. (1995). Spatial alignment across gaps: contributions of orientation and spatial scale. *Journal of the Optical Society of America A*, *12*, 2305–2317.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: a single-channel theory. *Vision Research*, *33*, 463–485.

Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology*, *14*, 570–572.

Wilcox, L. M., & Hess, R. F. (1996). Is the site of non-linear filtering stereopsis before or after the site of binocular combination? *Vision Research*, *36*, 391–400.

Wilson, H., & Bergen, J. (1979). A four-mechanism model for threshold spatial vision. *Vision Research*, *19*, 19–32.

Wilson, H. R., & Gelb, D. J. (1984). Modified line element theory for spatial frequency and width discrimination. *Journal of the Optical Society of America A*, *1*, 124–131.

Wilson, H. R., & Humanski, R. (1993). Spatial frequency adaptation andcontrast gain control. *Vision Research*, *33*, 1133–1149.

Wilson, H. R., & Kim, J. (1994). A model for motion coherence and transparency. *Visual Neuroscience*, *11*, 1205–1220.

Wilson, H. R., & Richards, W. A. (1992). Curvature separation and discrimination at texture boundaries. *Journal of the Optical Society of America A*, *9*, 1653–1662.

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research*, *26*, 453–469.

Wilson, H. R. (1993). Nonlinear processes in visual pattern discrimination. *Proceedings of the National Academy of Science USA*, *90*, 9785–9790.

Wilson, H. R. (1994). *Role of second-order motion signals in coherence and transparency*. *Higher order processing in the visual system*. (Ciba Foundation Symposium 184. New York: Wiley, 227–237.

Zenger, B., & Sagi, D. (1996). Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vision Research*, *36*, 2497–2513.